IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Ling, et al

Serial No: 09/883,848

Filed:

June 18, 2001

For:

MATE 7007 MCELLED Examiner: Angiogenesis-Modulating Compositions

and Uses

MAR 0 6 2002 Agtorney Docket⊗No.

CIBT-P01-119

1646

Art Unite AD

Not Assigned

COPY OF PAPERS ORIGINALLY FILED

CERTIFICATE OF MAILING UNDER 37 C.F.R. §1.8(a)

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as First Class Mail in an envelope addressed to Assistant Commissioner for Patents, Washington, D.C. 20231

February 26, 2002

Date

TRANSMITTAL OF SUBSTITUTE SEQUENCE LISTING IN COMPUTER REALESTORM 1600/2900

IN COMPLIANCE WITH 37 C.F.R. §1.825

Box Sequence Assistant Commissioner for Patents Washington, D.C. 20231

Sir:

This amendment is submitted in response to the Notice to Comply with Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Sequence Disclosures mailed on February 15, 2002; a copy is attached.

Transmitted herewith is a copy of a "Substitute Sequence Listing" in paper form for the above-identified patent application as required by 37 C.F.R. §1.825(a), and a copy of the "Substitute Sequence Listing" in computer readable form as required by 37 C.F.R. §§ 1.825(b).

As required by 37 C.F.R. § 1.825(b), Applicant's Agent hereby states that the contents of the Substitute "Sequence Listing" in paper form and in the computer readable form submitted herewith are the same and, as required by 37 C.F.R. § 1.825(a), also states that the submission includes no new matter.

Respectfully Submitted,

Date: February 26, 2002

Customer No: 28120 **Docketing Specialist** Ropes & Gray One International Place Boston, MA 02110

Phone: 617-951-7615 Fax: 617-951-7050

David P. Halstead

Registration No.: 44,735

COMMISSIONER FOR PATENTS UNITED STATES PATENT AND TRADEMARK OFFICE WASHINGTON, D.C. 2023!

www.uspto.gov ATTORNEY DOCKET NUMBER FIRST NAMED APPLICANT APPLICATION NUMBER FILING/RECEIPT &Leona E. Ling 06/18/200 CIBT-P01-119 09/883,848 MAR 0 6 20032 **CONFIRMATION NO. 9957 FORMALITIES LETTER** Ropes & Creav *OC00000007482459*

28120 **ROPES & GRAY** ONE INTERNATIONAL PLACE BOSTON, MA 02110-2624

Intellectual Property Dept. FEB 2 0 2002

COPY OF PAPERS ORIGINALLY FILED Date Mailed: 02/15/2002

NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES

Applicant is given TWO MONTHS FROM THE DATE OF THIS NOTICE within which to file the items indicated below to avoid abandonment. Extensions of time may be obtained under the provisions of 37 CFR 1.136(a)

· A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 C.F.R. 1.822 and/or 1.823, as indicated on the attached copy of the marked -up "Raw Sequence Listing." Applicant must provide a substitute computer readable form (CRF) copy of the "Sequence Listing" and a statement that the content of the sequence listing information recorded in computer readable form is identical to the written (on paper or compact disc) sequence listing and, where applicable, includes no new matter, as required by 37 CFR 1.821(e), 1.821(f), 1.821(g), 1.825(b), or 1.825(d).

For questions regarding compliance to these requirements, please contact:

- For Rules Interpretation, call (703) 308-4216
- To Purchase Patentin Software, call (703) 306-2600
- For Patentin Software Program Help, call (703) 306-4119 or e-mail at patin21help@uspto.gov or patin3help@uspto.gov

A copy of this notice <u>MUST</u> be returned with the reply.

Customer Service Center

Initial Patent Examination Division (703) 308-1202

PART 1 - ATTORNEY/APPLICANT COPY



MANLO DATE CANCELLED

MAR TO B 2002

UNITED STATES PATENT AND TRADEMARK OFFICE

IESOS .D.C. DOINHRAW VOQ.OTQU.WWW

APPLICATION NUMBER

FILING/RECEIPT DATE

United States Patent and Trademark Office

FIRST NAMED APPLICANT

ATTORNEY DOCKET NUMBER

09/883,848

06/18/2001

Leona E. Ling

CIBT-P01-119

CONFIRMATION NO. 9957

28120 ROPES & GRAY ONE INTERNATIONAL PLACE

BOSTON, MA 02110-2624

COPY OF PAPERS ORIGINALLY FILED

FORMALITIES LETTER

OC00000007482459

RECEIVED

MAR 1.5 2,352

Date Mailed: 02/15/2002

NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES

Applicant is given TWO MONTHS FROM THE DATE OF THIS NOTICE within which to file the items indicated below to avoid abandonment. Extensions of time may be obtained under the provisions of 37 CFR 1.136(a).

• A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 C.F.R. 1.822 and/or 1.823, as indicated on the attached copy of the marked -up "Raw Sequence Listing." Applicant must provide a substitute computer readable form (CRF) copy of the "Sequence Listing" and a statement that the content of the sequence listing information recorded in computer readable form is identical to the written (on paper or compact disc) sequence listing and, where applicable, includes no new matter, as required by 37 CFR 1.821(e), 1.821(f), 1.821(g), 1.825(b), or 1.825(d).

For questions regarding compliance to these requirements, please contact:

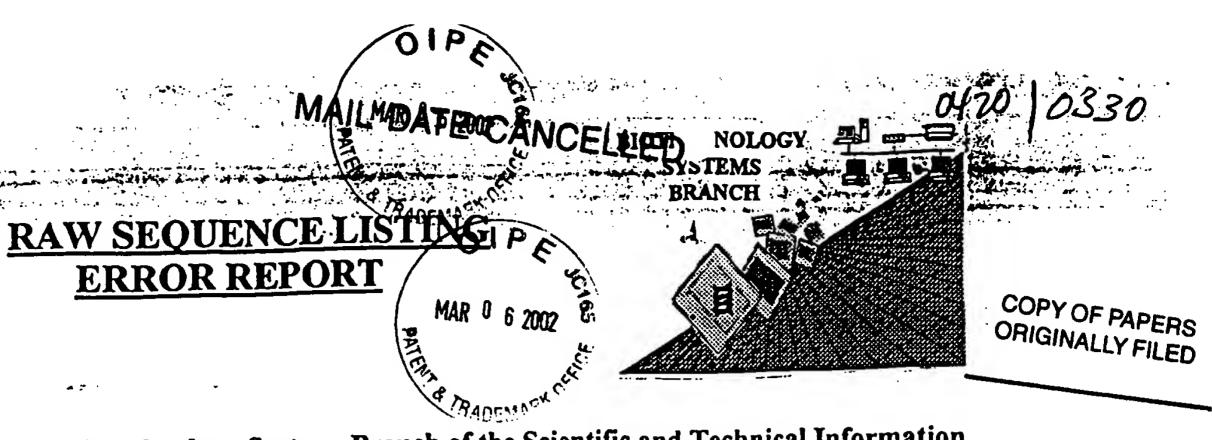
- For Rules Interpretation, call (703) 308-4216
- To Purchase Patentin Software, call (703) 306-2600
- For Patentin Software Program Help, call (703) 306-4119 or e-mail at patin21help@uspto.gov or patin3help@uspto.gov

A copy of this notice MUST be returned with the reply.

Customer Service Center

Initial Patent Examination Division (703) 308-1202

PART 2 - COPY TO BE RETURNED WITH RESPONSE



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/883,848

Source:

RECEIVED MAR 13 2002

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION, WITH A NOTICE TO COMPLY APPLICANT, WITH A NOTICE TO COMPLY APPLICANT. NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216. PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax) PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2Kcompliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address: http://www.uspto.gov/web/offices/pac/checker

	/200 21/0
ERROR DETECTED	SUGGESTED CORRECTION SERIAL NUMBER: 09/883, 848
ATTN: NEW RULES CASES	: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWAF
1Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
2Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.
3Misaligned Amino Numbering	The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
4Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
5Variable Length	Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
6PatentIn 2.0 "bug"	A "bug" in Patentin version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) Normally, Patentin would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
7Skipped Sequences (OLD RULES)	Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped
	Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to Include the skipped sequences.
8Skipped Sequences' (NEW RULES)	Sequence(s) missing. If Intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000
9 Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
10Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
11Use of <220>	Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
12PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of Patentin version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
13Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH - Biotechnology Systems Branch - 08/21/2001

OIPE

```
DATE: 10/12/2001
RAW SEQUENCE LISTING
                                        TIME: 12:14:38
PATENT APPLICATION: US/09/883,848
Input Set : A:\CIBT-P01-119 Seq List.txt
```

Does Not Comply Corrected Diskette Needed

```
Output Set: N:\CRF3\10122001\1883848.raw
      3 <110> APPLICANT: Ling, L.
              Sanicola-Nadel, M.
      6 <120> TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES
      8 <130> FILE REFERENCE: CIBT-P01-119
     10 <140> CURRENT APPLICATION NUMBER: 09/883,848
11 <141> CURRENT FILING DATE: 2001-09-24
     13 <150> PRIOR APPLICATION NUMBER: 60/211,919
     14 <151> PRIOR FILING DATE: 2000-06-16
     16 <160> NUMBER OF SEQ ID NOS: 48
     18 <170> SOFTWARE: PatentIn Ver. 2.1
```

ERRORED SEQUENCES

```
167 <210> SEQ ID NO: 6
    168 <211> LENGTH: 1425
    169 <212> TYPE: DNA
    170 <213> ORGANISM: Homo sapiens
    172 <400> SEQUENCE: 6
    173 atgctgctgc tggcgagatg tctgctgcta gtcctcgtct cctcgctgct ggtatgctcg 60
    174 ggactggcgt gcggaccggg cagggggttc gggaagagga ggcaccccaa aaagctgacc 120
    175 cctttagcct acaagcagtt tatccccaat gtggccgaga agaccctagg cgccagcgga 180
    176 aggtatgaag ggaagatete cagaaactee gagegattta aggaaeteae eeccaattae 240
    177 aaccccgaca tcatatttaa ggatgaagaa aacaccggag cggacaggct gatgactcag 300
    178 aggtgtaagg acaagttgaa cgctttggcc atctcggtga tgaaccagtg gccaggagtg 360
    179 aaactgcggg tgaccgaggg ctgggacgaa gatggccacc actcagagga gtctctgcac 420
    180 tacgagggcc gcgcagtgga catcaccacg tctgaccgcg accgcagcaa gtacggcatg 480
    181 ctggcccgcc tggcggtgga ggccggcttc gactgggtgt actacgagtc caaggcacat 540
    182 atccactgct cggtgaaagc agagaactcg gtggcggcca aatcgggagg ctgcttcccg 600
    183 ggctcggcca cggtgcacct ggagcagggc ggcaccaagc tggtgaagga cctgagcccc 660
    184 ggggaccgcg tgctggcggc ggacgaccag ggccggctgc tctacagcga cttcctcact 720
    185 ttectggace gegacgaegg egecaagaag gtettetaeg tgategagae gegggageeg 780
     186 cgcgagcgcc tgctgctcac cgccgcgcac ctgctctttg tggcgccgca caacgactcg 840.
     187 gccaccgggg agcccgaggc gtcctcgggc tcggggccgc cttccggggg cgcactgggg 900
     188 cctcgggcgc tgttcgccag ccgcgtgcgc ccgggccagc gcgtgtacgt ggtggccgag 960
     189 cgtgacgggg accgccggct cctgcccgcc gctgtgcaca gcgtgaccct aagcgaggag 1020
     190 gccgcgggcg cctacgcgcc gctcacggcc cagggcacca ttctcatcaa ccgggtgctg 1080
     191 gcctcgtgct acgcggtcat cgaggagcac agctgggcgc accgggcctt cgcgcccttc 1140
     192 cgcctggcgc acgcgctcct ggctgcactg gcgcccgcgc gcacggaccg cggcggggac 1200
     193 agcggcggcg gggaccgcgg gggcggcggc ggcagagtag ccctaaccgc tccaggtgct 1260
     194 gccgacgctc cgggtgcggg ggccaccgcg ggcatccact ggtactcgca gctgctctac 1320
     195 caaataggca cctggctcct ggacagcgag gccctgcacc cgctgggcat ggcggtcaag 1380
E--> 196 tccagdnina gccggggggc cgggggaggg gcgcgggagg gggcc
     728 <210> SEQ ID NO: 15
     729 <211> LENGTH: 475
     731 <213> ORGANISM: Homo sapiens
```

item 9 on Enor the



PATENT APPLICATION: US/09/883,848

DATE: 10/12/2001 TIME: 12:14:38

Input Set: A:\CIBT-P01-119 Seq List.txt
Output Set: N:\CRF3\10122001\I883848.raw

733 <400> SEQUENCE: 15 734 Met Leu Leu Leu Ala Arg Cys Leu Leu Val Leu Val Ser Ser Leu 737 Leu Val Cys Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly Lys 740 Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile 743 Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly 746 Lys Ile Ser Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr 747 65 749 Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg 752 Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile Ser 755 Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp 758 Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly Arg 761 Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly Met 762 145 764 Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu 767 Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala 770 Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu Glu 773 Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly Asp Arg Val 776 Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu Thr 777 225 779 Phe Leu Asp Arg Asp Asp Gly Ala Lys Lys Val Phe Tyr Val Ile Glu 782 Thr Arg Glu Pro Arg Glu Arg Leu Leu Leu Thr Ala Ala His Leu Leu 785 Phe Val Ala Pro His Asn Asp Ser Ala Thr Gly Glu Pro Glu Ala Ser 788 Ser Gly Ser Gly Pro Pro Ser Gly Gly Ala Leu Gly Pro Arg Ala Leu 791 Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val Ala Glu 792 305 794 Arg Asp Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser Val Thr 797 Leu Ser Glu Glu Ala Ala Gly Ala Tyr Ala Pro Leu Thr Ala Gln Gly 800 Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val Ile Glu 803 Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu Ala His

COPY OF PAPERS ORIGINALLY FILED

DATE: 10/12/2001

TIME: 12:14:38

95

175 —

125

PATENT APPLICATION: US/09/883,848 Input Set : A:\CIBT-P01-119 Seq List.txt Output Set: N:\CRF3\10122001\1883848.raw 806 Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Arg Gly Gly Asp 395 807 385 809 Ser Gly Gly Gly Asp Arg Gly Gly Gly Gly Arg Val Ala Leu Thr 410 405 812 Ala Pro Gly Ala Ala Asp Ala Pro Gly Ala Gly Ala Thr Ala Gly Ile 425 420 815 His Trp Tyr Ser Gln Leu Leu Tyr Gln Ile Gly Thr Trp Leu Leu Asp 440 435 E--> 818 Ser Glu Ala Leu His Pro Leu Gly Met Ala Val Lys Ser Ser Xaa Ser 816 455 450 819 821 Arg Gly Ala Gly Gly Gly Ala Arg Glu Gly Ala 475 470 822 465 1348 <210> SEQ ID NO: 23 1349 <211> LENGTH: (174) 1350 <212> TYPE: PRT 1351 <213> ORGANISM: Homo sapiens 1353 <400> SEQUENCE: 23 1354 Cys Gly Pro Gly Arg Val Val Gly Ser Arg Arg Pro Pro Arg Lys 10 5 1355 1357 Leu Val Pro Leu Ala Tyr Lys Gln Phe Ser Pro Asn Val Pro Glu Lys 30 25 1358 1360 Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly Lys Ile Ala Arg Ser Ser 45 40 35 1361 1363 Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe 55 50 1364 1366 Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg Leu Met Thr Gln Arg Cys 75 70 1367 65

1369 Lys Asp Arg Leu Asn Ser Leu Ala Ile Ser Val Met Asn Gln Trp Pro

1372 Gly Val Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His His

1375 Ser Glu Glu Ser Leu His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr

1378 Ser Asp Arg Asp Arg Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala Val

1381 Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu Ser Lys Ala His Val His

135

1384 Cys Ser Val Lys Ser Glu His Ser Ala Ala Lys Thr Gly Gly

150

120

85

165

100

115

130

RAW SEQUENCE LISTING

1370

1373

1376

1379

E--> 1385

1382 145

Use of n and/or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to insure a corresponding explanation is presented in the <220> to <223> fields of each sequence using n or Xaa.

170

90

155

105

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/883,848
DATE: 10/12/2001
TIME: 12:14:39

Input Set : A:\CIBT-P01-119 Seq List.txt
Output Set: N:\CRF3\10122001\I883848.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:196 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:6
L:818 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:15
L:1385 M:252 E: No. of Seq. differs, <211>LENGTH:Input:174 Found:175 SEQ:23
L:1698 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1701 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1704 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1707 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1711 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1714 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1720 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1723 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1726 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1729 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1729 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:1729 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26

Ţ

(L





COPY OF PAPERS

ORIGINALLY FILED

SEQUENCE LISTING

<110> Ling, L. Sanicola-Nadel, M.

<120> ANGIOGENESIS-MODULATING COMPOSITIONS AND USES

<130> CIBT-P01-119

<140> 09/883,848

<141> 2001-06-18

<150> 60/211,919

<151> 2000-06-16

<160> 48

<170> PatentIn Ver. 2.1

<210> 1

<211> 1277

<212> DNA

<213> Gallus gallus

<400> 1

atggtcgaaa tgctgctgtt gacaagaatt ctcttggtgg gcttcatctg cgctctttta 60 gtctcctctg ggctgacttg tggaccaggc aggggcattg gaaaaaggag gcaccccaaa 120 aagctgaccc cgttagccta taagcagttt attcccaatg tggcagagaa gaccctaggg 180 gccagtggaa gatatgaagg gaagatcaca agaaactccg agagatttaa agaactaacc 240 ccaaattaca accetgacat tattttaag gatgaagaga acaegggage tgacagaetg 300 atgactcagc gctgcaagga caagctgaat gccctggcga tctcggtgat gaaccagtgg 360 cccggggtga agctgcgggt gaccgagggc tgggacgagg atggccatca ctccgaggaa 420 tegetgeact acgagggteg egeegtggae ateaceaegt eggateggga eegeageaag 480 tacggaatgc tggcccgcct cgccgtcgag gccggcttcg actgggtcta ctacgagtcc 540 aaggcgcaca tccactgctc cgtcaaagca gaaaactcag tggcagcgaa atcaggaggc 600 tgcttccctg gctcagccac agtgcacctg gagcatggag gcaccaagct ggtgaaggac 660 ctgagccctg gggaccgcgt gctggctgct gacgcggacg gccggctgct ctacagtgac 720 ttcctcacct tcctcgaccg gatggacagc tcccgaaagc tcttctacgt catcgagacg 780 eggeageece gggeeegget getactgaeg geggeeeace tgetetttgt ggeeeeceag 840 cacaaccagt cggaggccac agggtccacc agtggccagg cgctcttcgc cagcaacgtg 900 aagcctggcc aacgtgtcta tgtgctgggc gagggcgggc agcagctgct gccggcgtct 960 gtccacageg teteattgeg ggaggaggeg teeggageet aegeeceaet cacegeecag 1020 ggcaccatcc tcatcaaccg ggtgttggcc tcctgctacg ccgtcatcga ggagcacagt 1080 tgggcccatt gggccttcgc accattccgc ttggctcagg ggctgctggc cgccctctgc 1140 ccagatgggg ccatccctac tgccgccacc accaccactg gcatccattg gtactcacgg 1200 ctcctctacc gcatcggcag ctgggtgctg gatggtgacg cgctgcatcc gctgggcatg 1260 1277 gtggcaccgg ccagctg

<210> 2

<211> 1190

<212> DNA

<213> Mus musculus

<400> 2

atggetetge eggeeagtet gttgeeetg tgetgettgg eactettgge actatetgee 60 cagagetgeg ggeeggeeg aggaeeggtt ggeeggegge gttatgtgeg eaageaactt 120

```
gtgcctctgc tatacaagca gtttgtgccc agtatgcccg agcggaccct gggcgcgagt 180
gggccagcgg aggggagggt aacaaggggg tcggagcgct tccgggacct cgtacccaac 240
tacaaccccg acataatctt caaggatgag gagaacagcg gcgcagaccg cctgatgaca 300
gagcgttgca aagagcgggt gaacgctcta gccatcgcgg tgatgaacat gtggcccgga 360
gtacgcctac gtgtgactga aggctgggac gaggacggcc accacgcaca ggattcactc 420
cactacgaag gccgtgcctt ggacatcacc acgtctgacc gtgaccgtaa taagtatggt 480
ttgttggcgc gcctagctgt ggaagccgga ttcgactggg tctactacga gtcccgcaac 540
cacatccacg tatcggtcaa agctgataac tcactggcgg tccgagccgg aggctgcttt 600
ccgggaaatg ccacggtgcg cttgcggagc ggcgaacgga aggggctgag ggaactacat 660
cgtggtgact gggtactggc cgctgatgca gcgggccgag tggtacccac gccagtgctg 720
ctcttcctgg accgggatct gcagcgccgc gcctcgttcg tggctgtgga gaccgagcgg 780
cctccgcgca aactgttgct cacaccctgg catctggtgt tcgctgctcg cgggccagcg 840
cetgetecag gtgaetttge accggtgtte gegegeeget taegtgetgg egaeteggtg 900
ctggctcccg gcggggacgc gctccagccg gcgcgcgtag cccgcgtggc gcgcgaggaa 960
gccgtgggcg tgttcgcacc gctcactgcg cacgggacgc tgctggtcaa cgacgtcctc 1020
gcctcctgct acgcggttct agagagtcac cagtgggccc accgcgcctt cgcccctttg 1080
cggctgctgc acgcgctcgg ggctctgctc cctgggggtg cagtccagcc gactggcatg 1140
cattggtact ctcgcctcct ttaccgcttg gccgaggagt taatgggctg
                                                                  1190
<210> 3
<211> 1281
<212> DNA
<213> Mus musculus
<400> 3
atgtctcccg cctggctccg gccccgactg cggttctgtc tgttcctgct gctgctgctt 60
ctggtgccgg cggcggggg ctgcgggccg ggccgggtgg tgggcagccg ccggaggccg 120
cctcgcaagc tcgtgcctct tgcctacaag cagttcagcc ccaacgtgcc ggagaagacc 180
ctgggcgcca gcgggcgcta cgaaggcaag atcgcgcgca gctctgagcg cttcaaagag 240
ctcaccccca actacaatcc cgacatcatc ttcaaggacg aggagaacac gggtgccgac 300
cgcctcatga cccagcgctg caaggaccgt ctgaactcac tggccatctc tgtcatgaac 360
cagtggcctg gtgtgaaact gcgggtgacc gaaggccggg atgaagatgg ccatcactca 420
gaggagtett tacactatga gggeegegeg gtggatatea eeaceteaga eegtgaeega 480
aataagtatg gactgctggc gcgcttagca gtggaggccg gcttcgactg ggtgtattac 540
gagtccaagg cccacgtgca ttgctctgtc aagtctgagc attcggccgc tgccaagaca 600
ggtggctgct ttcctgccgg agcccaggtg cgcctagaga acggggagcg tgtggccctg 660
tcagctgtaa agccaggaga ccgggtgctg gccatggggg aggatgggac ccccaccttc 720
agtgatgtgc ttattttcct ggaccgcgag ccaaaccggc tgagagcttt ccaggtcatc 780
gagactcagg atcetecgeg teggetggeg etcaegeetg eccaectget etteattgeg 840
gacaatcata cagaaccage ageceaette egggeeaeat ttgeeageea tgtgeaacca 900
ggccaatatg tgctggtatc aggggtacca ggcctccagc ctgctcgggt ggcagctgtc 960
tecaeceaeg tggeeettgg gteetatget eeteteaeaa ggeatgggae aettgtggtg 1020
gaggatgtgg tggcctcctg ctttgcagct gtggctgacc accatctggc tcagttggcc 1080
ttctggcccc tgcgactgtt tcccagtttg gcatggggca gctggacccc aagtgagggt 1140
gttcactcct accetcagat getctacege etggggegte tettgetaga agagageace 1200
ttccatccac tgggcatgtc tggggcagga agctgaaggg actctaacca ctgccctcct 1260
ggaactgctg tgcgtggatc c
                                                                  1281
<210> 4
<211> 1313
<212> DNA
<213> Mus musculus
```

atgctgctgc tgctggccag atgttttctg gtgatccttg cttcctcgct gctggtgtgc 60

<400> 4

```
cccgggctgg cctgtgggcc cggcaggggg tttggaaaga ggcggcaccc caaaaagctg 120
acceptitag cetacaagea gittatieee aacgiageeg agaagaeeet aggggeeage 180
ggcagatatg aagggaagat cacaagaaac tccgaacgat ttaaggaact cacccccaat 240
tacaaccccg acatcatatt taaggatgag gaaaacacgg gagcagaccg gctgatgact 300
cagaggtgca aagacaagtt aaatgccttg gccatctctg tgatgaacca gtggcctgga 360
gtgaggctgc gagtgaccga gggctgggat gaggacggcc atcattcaga ggagtctcta 420
cactatgagg gtcgagcagt ggacatcacc acgtccgacc gggaccgcag caagtacggc 480
atgctggctc gcctggctgt ggaagcaggt ttcgactggg tctactatga atccaaagct 540
cacatccact gttctgtgaa agcagagaac tccgtggcgg ccaaatccgg cggctgtttc 600
ccgggatccg ccaccgtgca cctggagcag ggcggcacca agctggtgaa ggacttacgt 660
cccggagacc gcgtgctggc ggctgacgac cagggccggc tgctgtacag cgacttcctc 720
accttcctgg accgcgacga aggcgccaag aaggtcttct acgtgatcga gacgctggag 780
ccgcgcgagc gcctgctgct caccgccgcg cacctgctct tcgtggcgcc gcacaacgac 840
teggggeeca egecegggee aagegegete tttgeeagee gegtgegeee egggeagege 900
gtgtacgtgg tggctgaacg cggcggggac cgccggctgc tgcccgccgc ggtgcacagc 960
gtgacgctgc gagaggagga ggcgggcgcg tacgcgccgc tcacggcgca cggcaccatt 1020
ctcatcaacc gggtgctcgc ctcgtgctac gctgtcatcg aggagcacag ctgggcacac 1080
egggeetteg egeettteeg eetggegeac gegetgetgg eegegetgge accegeeege 1140
acggacggcg ggggcggggg cagcatccct gcagcgcaat ctgcaacgga agcgaggggc 1200
gcggagccga ctgcgggcat ccactggtac tcgcagctgc tctaccacat tggcacctgg 1260
ctgttggaca gcgagaccat gcatcccttg ggaatggcgg tcaagtccag ctg
                                                                  1313
<210> 5
<211> 1256
<212> DNA
<213> Brachydanio rerio
<400> 5
atgcggcttt tgacgagagt gctgctggtg tctcttctca ctctgtcctt ggtggtgtcc 60
ggactggcct gcggtcctgg cagaggctac ggcagaagaa gacatccgaa gaagctgaca 120
cctctcgcct acaagcagtt catacctaat gtcgcggaga agaccttagg ggccagcggc 180
agatacgagg gcaagataac gcgcaattcg gagagattta aagaacttac tccaaattac 240
aatcccgaca ttatctttaa ggatgaggag aacacgggag cggacaggct catgacacag 300
agatgcaaag acaagctgaa ctcgctggcc atctctgtaa tgaaccactg gccaggggtt 360
aagctgcgtg tgacagaggg ctgggatgag gacggtcacc attttgaaga atcactccac 420
tacgagggaa gagctgttga tattaccacc tctgaccgag acaagagcaa atacgggaca 480
ctgtctcgcc tagctgtgga ggctggattt gactgggtct attacgagtc caaagcccac 540
atteattget etgteaaage agaaaatteg gttgetgega aatetggggg etgttteeea 600
ggttcggctc tggtctcgct ccaggacgga ggacagaagg ccgtgaagga cctgaacccc 660
ggagacaagg tgctggcggc agacagcgcg ggaaacctgg tgttcagcga cttcatcatg 720
ttcacagacc gagactccac gacgcgacgt gtgttttacg tcatagaaac gcaagaaccc 780
gttgaaaaga tcaccctcac cgccgctcac ctcctttttg tcctcgacaa ctcaacggaa 840
gatetecaea ceatgacege egegtatgee ageagtgtea gageeggaea aaaggtgatg 900
gttgttgatg atagcggtca gcttaaatct gtcatcgtgc agcggatata cacggaggag 960
cagcggggct cgttcgcacc agtgactgca catgggacca ttgtggtcga cagaatactg 1020
gcgtcctgtt acgccgtaat agaggaccag gggcttgcgc atttggcctt cgcgcccgcc 1080
aggetetatt attacgtgte atcattectg tececeaaaa etecageagt eggteeaatg 1140
cgactttaca acaggagggg gtccactggt actccaggct cctgtcatca aatgggaacg 1200
tggcttttgg acagcaacat gcttcatcct ttggggatgt cagtaaactc aagctg
                                                                  1256
```

```
<210> 6
<211> 1425
<212> DNA
<213> Homo sapiens
```

```
<221> modified_base
<222> (1387...1389)
<223> n=a, c, g, or t
<400> 6
atgctgctgc tggcgagatg tctgctgcta gtcctcgtct cctcgctgct ggtatgctcg 60
ggactggcgt gcggaccggg cagggggttc gggaagagga ggcaccccaa aaagctgacc 120
cctttagcct acaagcagtt tatccccaat gtggccgaga agaccctagg cgccagcgga 180
aggtatgaag ggaagatete cagaaactee gagegattta aggaacteae eeccaattae 240
aaccccgaca tcatatttaa ggatgaagaa aacaccggag cggacaggct gatgactcag 300
aggtgtaagg acaagttgaa cgctttggcc atctcggtga tgaaccagtg gccaggagtg 360
aaactgcggg tgaccgaggg ctgggacgaa gatggccacc actcagagga gtctctgcac 420
tacgagggcc gcgcagtgga catcaccacg tctgaccgcg accgcagcaa gtacggcatg 480
ctggcccgcc tggcggtgga ggccggcttc gactgggtgt actacgagtc caaggcacat 540
atccactgct cggtgaaagc agagaactcg gtggcggcca aatcgggagg ctgcttcccg 600
ggctcggcca cggtgcacct ggagcagggc ggcaccaagc tggtgaagga cctgagcccc 660
ggggaccgcg tgctggcggc ggacgaccag ggccggctgc tctacagcga cttcctcact 720
ttcctggacc gcgacgacgg cgccaagaag gtcttctacg tgatcgagac gcgggagccg 780
cgcgagcgcc tgctgctcac cgccgcgcac ctgctctttg tggcgccgca caacgactcg 840
gccaccgggg agcccgaggc gtcctcgggc tcggggccgc cttccggggg cgcactgggg 900
cctcgggcgc tgttcgccag ccgcgtgcgc ccgggccagc gcgtgtacgt ggtggccgag 960
cgtgacgggg accgccggct cctgcccgcc gctgtgcaca gcgtgaccct aagcgaggag 1020
geegegggeg cetaegegee geteaeggee eagggeaeca tteteateaa eegggtgetg 1080
gcctcgtgct acgcggtcat cgaggagcac agctgggcgc accgggcctt cgcgcccttc 1140
cgcctggcgc acgcgctcct ggctgcactg gcgcccgcgc gcacggaccg cggcggggac 1200
ageggeggeg gggacegegg gggeggegge ggeagagtag ceetaacege teeaggtget 1260
geegaegete egggtgeggg ggeeaeegeg ggeateeaet ggtaetegea getgetetae 1320
caaataggca cctggctcct ggacagcgag gccctgcacc cgctgggcat ggcggtcaag 1380
                                                                  1425
tccagcnnna gccggggggc cggggggggg gcgcgggagg gggcc
<210> 7
<211> 1622
<212> DNA
<213> Homo sapiens
<400> 7
catcagecea ecaggagace tegecegeeg eteceeeggg eteceeggee atgteteeeg 60
cccggctccg gccccgactg cacttctgcc tggtcctgtt gctgctgctg gtggtgcccg 120
cggcatgggg ctgcgggccg ggtcgggtgg tgggcagccg ccggcgaccg ccacgcaaac 180
tcgtgccgct cgcctacaag cagttcagcc ccaatgtgcc cgagaagacc ctgggcgcca 240
geggaegeta tgaaggeaag ategetegea geteegageg etteaaggag eteaeceeca 300
attacaatcc agacatcatc ttcaaggacg aggagaacac aggcgccgac cgcctcatga 360
cccagcgctg caaggaccgc ctgaactcgc tggctatctc ggtgatgaac cagtggcccg 420
gtgtgaaget gegggtgaee gagggetggg aegaggaegg ceaecaetea gaggagteee 480
tgcattatga gggccgcgcg gtggacatca ccacatcaga ccgcgaccgc aataagtatg 540
gactgctggc gcgcttggca gtggaggccg gctttgactg ggtgtattac gagtcaaagg 600
cccacgtgca ttgctccgtc aagtccgagc actcggccgc agccaagacg ggcggctgct 660
tecetgeegg ageceaggta egeetggaga gtggggegeg tgtggeettg teageegtga 720
ggccgggaga ccgtgtgctg gccatggggg aggatgggag ccccaccttc agcgatgtgc 780
tcattttcct ggaccgcgag ccccacaggc tgagagcctt ccaggtcatc gagactcagg 840
acceceacg eggetggea eteacaceg eteacetget etttaegget gacaateaca 900
cggagccggc agcccgcttc cgggccacat ttgccagcca cgtgcagcct ggccagtacg 960
tgctggtggc tggggtgcca ggcctgcagc ctgcccgcgt ggcagctgtc tctacacacg 1020
tggccctcgg ggcctacgcc ccgctcacaa agcatgggac actggtggtg gaggatgtgg 1080
tggcatcctg cttcgcggcc gtggctgacc accacctggc tcagttggcc ttctggcccc 1140
tgagactctt tcacagcttg gcatggggca gctggacccc gggggagggt gtgcattggt 1200
```

```
accccaget getetacege etggggegte teetgetaga agagggeage tteeacccae 1260
tgggcatgtc cggggcaggg agctgaaagg actccaccgc tgccctcctg gaactgctgt 1320
actgggtcca gaagcctctc agccaggagg gagctggccc tggaagggac ctgagctggg 1380
ggacactggc tcctgccatc tcctctgcca tgaagataca ccattgagac ttgactgggc 1440
aacaccagcg tececeaece gegtegtggt gtagteatag agetgeaage tgagetggeg 1500
aggggatggt tgttgacccc tctctcctag agaccttgag gctggcacgg cgactcccaa 1560
ctcagcctgc tctcactacg agttttcata ctctgcctcc cccattggga gggcccattc 1620
                                                                  1622
CC
<210> 8
<211> 1191
<212> DNA
<213> Homo sapiens
<400> 8
atggctctcc tgaccaatct actgcccttg tgctgcttgg cacttctggc gctgccagcc 60
cagagetgeg ggeegggeeg ggggeeggtt ggeeggegee getatgegeg caageagete 120
gtgccgctac tctacaagca atttgtgccc ggcgtgccag agcggaccct gggcgccagt 180
gggccagcgg aggggagggt ggcaaggggc tccgagcgct tccggggacct cgtgcccaac 240
tacaaccccg acatcatctt caaggatgag gagaacagtg gagccgaccg cctgatgacc 300
gagcgttgca aggagaggt gaacgctttg gccattgccg tgatgaacat gtggcccgga 360
gtgcgcctac gagtgactga gggctgggac gaggacggcc accacgctca ggattcactc 420
cactacgaag gccgtgcttt ggacatcact acgtctgacc gcgaccgcaa caagtatggg 480
ttgctggcgc gcctcgcagt ggaagccggc ttcgactggg tctactacga gtcccgcaac 540
cacgtccacg tgtcggtcaa agctgataac tcactggcgg tccgggcggg cggctgcttt 600
ccgggaaatg caactgtgcg cctgtggagc ggcgagcgga aagggctgcg ggaactgcac 660
cgcggagact gggttttggc ggccgatgcg tcaggccggg tggtgcccac gccggtgctg 720
ctcttcctgg accgggactt gcagcgccgg gcttcatttg tggctgtgga gaccgagtgg 780
cctccacgca aactgttgct cacgccctgg cacctggtgt ttgccgctcg agggccggcg 840
cccgcgccag gcgactttgc accggtgttc gcgcgccggc tacgcgctgg ggactcggtg 900
ctggcgcccg gcggggatgc gcttcggcca gcgcgcgtgg cccgtgtggc gcgggaggaa 960
geegtgggeg tgttegegee geteacegeg eaegggaege tgetggtgaa egatgteetg 1020
gcctcttgct acgcggttct ggagagtcac cagtgggcgc accgcgcttt tgcccccttg 1080
agactgctgc acgcgctagg ggcgctgctc cccggcgggg ccgtccagcc gactggcatg 1140
                                                                  1191
cattggtact ctcggctcct ctaccgctta gcggaggagc tactgggctg a
<210> 9
<211> 1251
<212> DNA
<213> Brachydanio rerio
<400> 9
atggacgtaa ggctgcatct gaagcaattt gctttactgt gttttatcag cttgcttctg 60
acgccttgtg gattagcctg tggtcctggt agaggttatg gaaaacgaag acacccaaag 120
aaattaaccc cgttggctta caagcaattc atccccaacg ttgctgagaa aacgcttgga 180
gccagcggca aatacgaagg caaaatcaca aggaattcag agagatttaa agagctgatt 240
ccgaattata atcccgatat catctttaag gacgaggaaa acacaaacgc tgacaggctg 300
atgaccaagc gctgtaagga caagttaaat tcgttggcca tatccgtcat gaaccactgg 360
cccggcgtga aactgcgcgt cactgaaggc tgggatgagg atggtcacca tttagaagaa 420
tctttgcact atgagggacg ggcagtggac atcactacct cagacaggga taaaagcaag 480
tatgggatgc tatccaggct tgcagtggag gcaggattcg actgggtcta ttatgaatct 540
aaagcccaca tacactgctc tgtcaaagca gaaaattcag tggctgctaa atcaggagga 600
tgttttcctg ggtctgggac ggtgacactt ggtgatggga cgaggaaacc catcaaagat 660
cttaaagtgg gcgaccgggt tttggctgca gacgagaagg gaaatgtctt aataagcgac 720
```

tttattatgt ttatagacca cgatccgaca acgagaaggc aattcatcgt catcgagacg 780

tcagaacctt tcaccaagct caccctcact gccgcgcacc tagttttcgt tggaaactct 840 tcagcagctt cgggtataac agcaacattt gccagcaacg tgaagcctgg agatacagtt 900 ttagtgtggg aagacacatg cgagagcctc aagagcgtta cagtgaaaag gatttacact 960 gaggagcacg agggctcttt tgcgccagtc accgcgcacg gaaccataat agtggatcag 1020 gtgttggcat cgtgctacgc ggtcattgag aaccacaaat gggcacattg ggcttttgcg 1080 ccggtcaggt tgtgtcacaa gctgatgacg tggcttttc cggctcgtga atcaaacgtc 1140 aattttcagg aggatggtat ccactggtac tcaaatatgc tgtttcacat cggctcttgg 1200 ctgctggaca gagactcttt ccatccactc gggattttac acttaagttg a 1251

<210> 10

<211> 425

<212> PRT

<213> Gallus gallus

<400> 10

Met Val Glu Met Leu Leu Thr Arg Ile Leu Leu Val Gly Phe Ile 1 5 10 15

Cys Ala Leu Leu Val Ser Ser Gly Leu Thr Cys Gly Pro Gly Arg Gly
20 25 30

Ile Gly Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys
35 40 45

Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg
50 55 60

Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr 65 70 75 80

Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly
85 90 95

Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu 100 105 110

Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr 115 120 125

Glu Gly Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr 130 135 140

Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys
145 150 155 160

Tyr Gly Met Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val 165 170 175

Tyr Tyr Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn 180 185 190

Ser Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val

His Leu Glu His Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly 210 215 220

Asp Arg Val Leu Ala Ala Asp Ala Asp Gly Arg Leu Leu Tyr Ser Asp 225 230 235 240

Phe Leu Thr Phe Leu Asp Arg Met Asp Ser Ser Arg Lys Leu Phe Tyr 245 250 255

Val Ile Glu Thr Arg Gln Pro Arg Ala Arg Leu Leu Leu Thr Ala Ala 260 265 270

His Leu Leu Phe Val Ala Pro Gln His Asn Gln Ser Glu Ala Thr Gly 275 280 285

Ser Thr Ser Gly Gln Ala Leu Phe Ala Ser Asn Val Lys Pro Gly Gln
290 295 300

Arg Val Tyr Val Leu Gly Glu Gly Gly Gln Gln Leu Leu Pro Ala Ser 305 310 315 320

Val His Ser Val Ser Leu Arg Glu Glu Ala Ser Gly Ala Tyr Ala Pro 325 330 335

Leu Thr Ala Gln Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys 340 345 350

Tyr Ala Val Ile Glu Glu His Ser Trp Ala His Trp Ala Phe Ala Pro 355 360 365

Phe Arg Leu Ala Gln Gly Leu Leu Ala Ala Leu Cys Pro Asp Gly Ala 370 375 380

Ile Pro Thr Ala Ala Thr Thr Thr Thr Gly Ile His Trp Tyr Ser Arg
385 390 395 400

Leu Leu Tyr Arg Ile Gly Ser Trp Val Leu Asp Gly Asp Ala Leu His
405 410 415

Pro Leu Gly Met Val Ala Pro Ala Ser 420 425

<210> 11

<211> 396

<212> PRT

<213> Mus musculus

<400> 11

Met Ala Leu Pro Ala Ser Leu Leu Pro Leu Cys Cys Leu Ala Leu Leu 1 5 10 15

Ala Leu Ser Ala Gln Ser Cys Gly Pro Gly Arg Gly Pro Val Gly Arg
20 25 30

Arg Arg Tyr Val Arg Lys Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe 35 40 45

Val Pro Ser Met Pro Glu Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu

60

Gly Arg Val Thr Arg Gly Ser Glu Arg Phe Arg Asp Leu Val Pro Asn 65 70 75 80

55

- Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp 90 95
- Arg Leu Met Thr Glu Arg Cys Lys Glu Arg Val Asn Ala Leu Ala Ile 100 105 110
- Ala Val Met Asn Met Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly
 115 120 125
- Trp Asp Glu Asp Gly His His Ala Gln Asp Ser Leu His Tyr Glu Gly
 130 135 140
- Arg Ala Leu Asp Ile Thr Thr Ser Asp Arg Asp Arg Asn Lys Tyr Gly
 145 150 155 160
- Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr
 165 170 175
- Glu Ser Arg Asn His Ile His Val Ser Val Lys Ala Asp Asn Ser Leu 180 185 190
- Ala Val Arg Ala Gly Gly Cys Phe Pro Gly Asn Ala Thr Val Arg Leu 195 200 205
- Arg Ser Gly Glu Arg Lys Gly Leu Arg Glu Leu His Arg Gly Asp Trp 210 215 220
- Val Leu Ala Asp Ala Ala Gly Arg Val Val Pro Thr Pro Val Leu 225 230 235 240
- Leu Phe Leu Asp Arg Asp Leu Gln Arg Arg Ala Ser Phe Val Ala Val 245 250 255
- Glu Thr Glu Arg Pro Pro Arg Lys Leu Leu Leu Thr Pro Trp His Leu 260 265 270
- Val Phe Ala Arg Gly Pro Ala Pro Ala Pro Gly Asp Phe Ala Pro 275 280 285
- Val Phe Ala Arg Arg Leu Arg Ala Gly Asp Ser Val Leu Ala Pro Gly 290 295 300
- Gly Asp Ala Leu Gln Pro Ala Arg Val Ala Arg Val Ala Arg Glu Glu 305 310 315 320
- Ala Val Gly Val Phe Ala Pro Leu Thr Ala His Gly Thr Leu Leu Val 325 330 335
- Asn Asp Val Leu Ala Ser Cys Tyr Ala Val Leu Glu Ser His Gln Trp 340 345 350
- Ala His Arg Ala Phe Ala Pro Leu Arg Leu Leu His Ala Leu Gly Ala

355 360 365

Leu Leu Pro Gly Gly Ala Val Gln Pro Thr Gly Met His Trp Tyr Ser 370 380

Arg Leu Leu Tyr Arg Leu Ala Glu Glu Leu Met Gly 395

<210> 12

<211> 411

<212> PRT

<213> Mus musculus

<400> 12

Met Ser Pro Ala Trp Leu Arg Pro Arg Leu Arg Phe Cys Leu Phe Leu
1 5 10 15

Leu Leu Leu Leu Val Pro Ala Ala Arg Gly Cys Gly Pro Gly Arg
20 25 30

Val Val Gly Ser Arg Arg Arg Pro Pro Arg Lys Leu Val Pro Leu Ala 35 40 45

Tyr Lys Gln Phe Ser Pro Asn Val Pro Glu Lys Thr Leu Gly Ala Ser 50 55 60

Gly Arg Tyr Glu Gly Lys Ile Ala Arg Ser Ser Glu Arg Phe Lys Glu 65 70 75 80

Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn 85 90 95

Thr Gly Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Arg Leu Asn 100 105 110

Ser Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg 115 120 125

Val Thr Glu Gly Arg Asp Glu Asp Gly His His Ser Glu Glu Ser Leu 130 135 140

His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg 145 150 150

Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp 165 170 175

Trp Val Tyr Tyr Glu Ser Lys Ala His Val His Cys Ser Val Lys Ser 180 185 190

Glu His Ser Ala Ala Ala Lys Thr Gly Gly Cys Phe Pro Ala Gly Ala 195 200 205

Gln Val Arg Leu Glu Asn Gly Glu Arg Val Ala Leu Ser Ala Val Lys 210 215 220 Pro Gly Asp Arg Val Leu Ala Met Gly Glu Asp Gly Thr Pro Thr Phe 225 230 235 240

Ser Asp Val Leu Ile Phe Leu Asp Arg Glu Pro Asn Arg Leu Arg Ala 245 250 255

Phe Gln Val Ile Glu Thr Gln Asp Pro Pro Arg Arg Leu Ala Leu Thr 260 265 270

Pro Ala His Leu Leu Phe Ile Ala Asp Asn His Thr Glu Pro Ala Ala 275 280 285

His Phe Arg Ala Thr Phe Ala Ser His Val Gln Pro Gly Gln Tyr Val 290 295 300

Leu Val Ser Gly Val Pro Gly Leu Gln Pro Ala Arg Val Ala Ala Val 305 310 315 320

Ser Thr His Val Ala Leu Gly Ser Tyr Ala Pro Leu Thr Arg His Gly 325 330 335

Thr Leu Val Val Glu Asp Val Val Ala Ser Cys Phe Ala Ala Val Ala 340 345 350

Asp His His Leu Ala Gln Leu Ala Phe Trp Pro Leu Arg Leu Phe Pro 355 360 365

Ser Leu Ala Trp Gly Ser Trp Thr Pro Ser Glu Gly Val His Ser Tyr 370 375 380

Pro Gln Met Leu Tyr Arg Leu Gly Arg Leu Leu Leu Glu Glu Ser Thr 385 390 395 400

Phe His Pro Leu Gly Met Ser Gly Ala Gly Ser 405 410

<210> 13

<211> 437

<212> PRT

<213> Mus musculus

<400> 13

Met Leu Leu Leu Ala Arg Cys Phe Leu Val Ile Leu Ala Ser Ser 1 5 10 15

Leu Leu Val Cys Pro Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly 20 25 30

Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe 35 40 45

Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu 50 60

Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn

Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp 90 95

Arg Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile 100 105 110

Ser Val Met Asn Gln Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly
115 120 125

Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly
130 135 140

Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly 145 150 155 160

Met Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr 165 170 175

Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val 180 185 190

Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu 195 200 205

Glu Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Arg Pro Gly Asp Arg 210 215 220

Val Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu 225 230 235 240

Thr Phe Leu Asp Arg Asp Glu Gly Ala Lys Lys Val Phe Tyr Val Ile 245 250 255

Glu Thr Leu Glu Pro Arg Glu Arg Leu Leu Leu Thr Ala Ala His Leu 260 265 270

Leu Phe Val Ala Pro His Asn Asp Ser Gly Pro Thr Pro Gly Pro Ser 275 280 285

Ala Leu Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val 290 295 300

Ala Glu Arg Gly Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser 305 310 315 320

Val Thr Leu Arg Glu Glu Glu Ala Gly Ala Tyr Ala Pro Leu Thr Ala 325 330 335

His Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val 340 345 350

Ile Glu Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu 355 360 365

Ala His Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Gly Gly

370 375 380

Gly Gly Ser Ile Pro Ala Ala Gln Ser Ala Thr Glu Ala Arg Gly 385 390 395 400

Ala Glu Pro Thr Ala Gly Ile His Trp Tyr Ser Gln Leu Leu Tyr His
405 410 415

Ile Gly Thr Trp Leu Leu Asp Ser Glu Thr Met His Pro Leu Gly Met 420 425 430

Ala Val Lys Ser Ser 435

<210> 14

<211> 418

<212> PRT

<213> Brachydanio rerio

<400> 14

Met Arg Leu Leu Thr Arg Val Leu Leu Val Ser Leu Leu Thr Leu Ser 1 10 15

Leu Val Val Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Tyr Gly Arg
20 25 30

Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile 35 40 45

Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly 50 55 60

Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr 65 70 75 80

Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg 85 90 95

Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ser Leu Ala Ile Ser 100 105 110

Val Met Asn His Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp 115 120 125

Asp Glu Asp Gly His His Phe Glu Glu Ser Leu His Tyr Glu Gly Arg
130 135 140

Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Lys Ser Lys Tyr Gly Thr 145 150 155 160

Leu Ser Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu 165 170 175

Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala 180 185 190 Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Leu Val Ser Leu Gln
195 200 205

Asp Gly Gln Lys Ala Val Lys Asp Leu Asn Pro Gly Asp Lys Val 210 215 220

Leu Ala Ala Asp Ser Ala Gly Asn Leu Val Phe Ser Asp Phe Ile Met 225 230 235 240

Phe Thr Asp Arg Asp Ser Thr Thr Arg Arg Val Phe Tyr Val Ile Glu 245 250 255

Thr Gln Glu Pro Val Glu Lys Ile Thr Leu Thr Ala Ala His Leu Leu 260 265 270

Phe Val Leu Asp Asn Ser Thr Glu Asp Leu His Thr Met Thr Ala Ala 275 280 285

Tyr Ala Ser Ser Val Arg Ala Gly Gln Lys Val Met Val Val Asp Asp 290 295 300

Ser Gly Gln Leu Lys Ser Val Ile Val Gln Arg Ile Tyr Thr Glu Glu 305 310 315 320

Gln Arg Gly Ser Phe Ala Pro Val Thr Ala His Gly Thr Ile Val Val
325 330 335

Asp Arg Ile Leu Ala Ser Cys Tyr Ala Val Ile Glu Asp Gln Gly Leu 340 345 350

Ala His Leu Ala Phe Ala Pro Ala Arg Leu Tyr Tyr Tyr Val Ser Ser 355 360 365

Phe Leu Ser Pro Lys Thr Pro Ala Val Gly Pro Met Arg Leu Tyr Asn 370 375 380

Arg Arg Gly Ser Thr Gly Thr Pro Gly Ser Cys His Gln Met Gly Thr 385 390 395 400

Trp Leu Leu Asp Ser Asn Met Leu His Pro Leu Gly Met Ser Val Asn 405 410 415

Ser Ser

<210> 15

<211> 475

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (463)

<223> Xaa=unknown amino acid residue

<400> 15

- Met Leu Leu Ala Arg Cys Leu Leu Leu Val Leu Val Ser Ser Leu
 1 5 10 15
- Leu Val Cys Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly Lys
 20 25 30
- Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile 35 40 45
- Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly 50 55 60
- Lys Ile Ser Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr 65 70 75 80
- Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg
 85 90 95
- Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile Ser 100 105 110
- Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp
 115 120 125
- Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly Arg 130 135 140
- Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly Met 145 150 155 160
- Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu 165 170 175
- Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala 180 185 190
- Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu Glu 195 200 205
- Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly Asp Arg Val 210 215 220
- Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu Thr 225 230 235 240
- Phe Leu Asp Arg Asp Asp Gly Ala Lys Lys Val Phe Tyr Val Ile Glu 245 250 255
- Thr Arg Glu Pro Arg Glu Arg Leu Leu Leu Thr Ala Ala His Leu Leu 260 265 270
- Phe Val Ala Pro His Asn Asp Ser Ala Thr Gly Glu Pro Glu Ala Ser 275 280 285
- Ser Gly Ser Gly Pro Pro Ser Gly Gly Ala Leu Gly Pro Arg Ala Leu 290 295 300

Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val Ala Glu 305 310 320

Arg Asp Gly Asp Arg Leu Leu Pro Ala Ala Val His Ser Val Thr 325 330 335

Leu Ser Glu Glu Ala Ala Gly Ala Tyr Ala Pro Leu Thr Ala Gln Gly 340 345 350

Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val Ile Glu 355 360 365

Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu Ala His 370 380

Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Arg Gly Gly Asp 385 390 395 400

Ser Gly Gly Asp Arg Gly Gly Gly Gly Arg Val Ala Leu Thr 405 410 415

Ala Pro Gly Ala Ala Asp Ala Pro Gly Ala Gly Ala Thr Ala Gly Ile 420 425 430

His Trp Tyr Ser Gln Leu Leu Tyr Gln Ile Gly Thr Trp Leu Leu Asp
435
440
445

Ser Glu Ala Leu His Pro Leu Gly Met Ala Val Lys Ser Ser Xaa Ser 450 455 460

Arg Gly Ala Gly Gly Gly Ala Arg Glu Gly Ala
465 470 475

<210> 16

<211> 411

<212> PRT

<213> Homo sapiens

<400> 16

Met Ser Pro Ala Arg Leu Arg Pro Arg Leu His Phe Cys Leu Val Leu 1 5 10 15

Leu Leu Leu Val Val Pro Ala Ala Trp Gly Cys Gly Pro Gly Arg
20 25 30

Val Val Gly Ser Arg Arg Pro Pro Arg Lys Leu Val Pro Leu Ala 35 40 45

Tyr Lys Gln Phe Ser Pro Asn Val Pro Glu Lys Thr Leu Gly Ala Ser 50 55 60

Gly Arg Tyr Glu Gly Lys Ile Ala Arg Ser Ser Glu Arg Phe Lys Glu 65 70 75 80

Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn 85 90 95

Thr Gly Ala Asp Arg Leu Met Thr Gln Arg Cys Lys Asp Arg Leu Asn 100 105 110

Ser Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg 115 120 125

Val Thr Glu Gly Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu 130 135 140

His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg 145 150 150

Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp 165 170 175

Trp Val Tyr Tyr Glu Ser Lys Ala His Val His Cys Ser Val Lys Ser 180 185 190

Glu His Ser Ala Ala Ala Lys Thr Gly Gly Cys Phe Pro Ala Gly Ala 195 200 205

Gln Val Arg Leu Glu Ser Gly Ala Arg Val Ala Leu Ser Ala Val Arg 210 215 220

Pro Gly Asp Arg Val Leu Ala Met Gly Glu Asp Gly Ser Pro Thr Phe 225 230 235 240

Ser Asp Val Leu Ile Phe Leu Asp Arg Glu Pro His Arg Leu Arg Ala 245 250 255

Phe Gln Val Ile Glu Thr Gln Asp Pro Pro Arg Arg Leu Ala Leu Thr 260 265 270

Pro Ala His Leu Leu Phe Thr Ala Asp Asn His Thr Glu Pro Ala Ala 275 280 285

Arg Phe Arg Ala Thr Phe Ala Ser His Val Gln Pro Gly Gln Tyr Val 290 295 300

Leu Val Ala Gly Val Pro Gly Leu Gln Pro Ala Arg Val Ala Ala Val 305 310 315 320

Ser Thr His Val Ala Leu Gly Ala Tyr Ala Pro Leu Thr Lys His Gly 325 330 335

Thr Leu Val Val Glu Asp Val Val Ala Ser Cys Phe Ala Ala Val Ala 340 345 350

Asp His His Leu Ala Gln Leu Ala Phe Trp Pro Leu Arg Leu Phe His 355 360 365

Ser Leu Ala Trp Gly Ser Trp Thr Pro Gly Glu Gly Val His Trp Tyr 370 380

Pro Gln Leu Leu Tyr Arg Leu Gly Arg Leu Leu Leu Glu Glu Gly Ser

Phe His Pro Leu Gly Met Ser Gly Ala Gly Ser 405 410

<210> 17

<211> 396

<212> PRT

<213> Homo sapiens

<400> 17

Met Ala Leu Leu Thr Asn Leu Leu Pro Leu Cys Cys Leu Ala Leu Leu 1 1 5 15

Ala Leu Pro Ala Gln Ser Cys Gly Pro Gly Arg Gly Pro Val Gly Arg
20 25 30

Arg Arg Tyr Ala Arg Lys Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe
35 40 45

Val Pro Gly Val Pro Glu Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu
50 55 60

Gly Arg Val Ala Arg Gly Ser Glu Arg Phe Arg Asp Leu Val Pro Asn 65 70 75 80

Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp 90 95

Arg Leu Met Thr Glu Arg Cys Lys Glu Arg Val Asn Ala Leu Ala Ile 100 105 110

Ala Val Met Asn Met Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly
115 120 125

Trp Asp Glu Asp Gly His His Ala Gln Asp Ser Leu His Tyr Glu Gly
130 135 140

Arg Ala Leu Asp Ile Thr Thr Ser Asp Arg Asp Arg Asn Lys Tyr Gly
145 150 155 160

Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr
165 170 175

Glu Ser Arg Asn His Val His Val Ser Val Lys Ala Asp Asn Ser Leu 180 185 190

Ala Val Arg Ala Gly Gly Cys Phe Pro Gly Asn Ala Thr Val Arg Leu 195 200 205

Trp Ser Gly Glu Arg Lys Gly Leu Arg Glu Leu His Arg Gly Asp Trp 210 215 220

Val Leu Ala Ala Asp Ala Ser Gly Arg Val Val Pro Thr Pro Val Leu

225	230	235	240
44			

Leu Phe Leu Asp Arg Asp Leu Gln Arg Arg Ala Ser Phe Val Ala Val
245 250 255

Glu Thr Glu Trp Pro Pro Arg Lys Leu Leu Leu Thr Pro Trp His Leu 260 270

Val Phe Ala Arg Gly Pro Ala Pro Ala Pro Gly Asp Phe Ala Pro 275 280 285

Val Phe Ala Arg Arg Leu Arg Ala Gly Asp Ser Val Leu Ala Pro Gly 290 295 300

Gly Asp Ala Leu Arg Pro Ala Arg Val Ala Arg Val Ala Arg Glu Glu 305 310 315 320

Ala Val Gly Val Phe Ala Pro Leu Thr Ala His Gly Thr Leu Leu Val
325 330 335

Asn Asp Val Leu Ala Ser Cys Tyr Ala Val Leu Glu Ser His Gln Trp
340 345 350

Ala His Arg Ala Phe Ala Pro Leu Arg Leu Leu His Ala Leu Gly Ala 355 360 365

Leu Leu Pro Gly Gly Ala Val Gln Pro Thr Gly Met His Trp Tyr Ser 370 375 380

Arg Leu Leu Tyr Arg Leu Ala Glu Glu Leu Leu Gly 385 390 395

<210> 18

<211> 416

<212> PRT

<213> Brachydanio rerio

<400> 18

Met Asp Val Arg Leu His Leu Lys Gln Phe Ala Leu Leu Cys Phe Ile 1 5 10 15

Ser Leu Leu Thr Pro Cys Gly Leu Ala Cys Gly Pro Gly Arg Gly
20 25 30

Tyr Gly Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys
35 40 45

Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Lys
50 55 60

Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Ile
65 70 75 80

Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Asn 85 90 95

- Ala Asp Arg Leu Met Thr Lys Arg Cys Lys Asp Lys Leu Asn Ser Leu 100 105 110
- Ala Ile Ser Val Met Asn His Trp Pro Gly Val Lys Leu Arg Val Thr 115 120 125
- Glu Gly Trp Asp Glu Asp Gly His His Leu Glu Glu Ser Leu His Tyr 130 135 140
- Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Lys Ser Lys 145 150 155 160
- Tyr Gly Met Leu Ser Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val 165 170 175
- Tyr Tyr Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn 180 185 190
- Ser Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Gly Thr Val 195 200 205
- Thr Leu Gly Asp Gly Thr Arg Lys Pro Ile Lys Asp Leu Lys Val Gly 210 215 220
- Asp Arg Val Leu Ala Ala Asp Glu Lys Gly Asn Val Leu Ile Ser Asp 225 230 235 240
- Phe Ile Met Phe Ile Asp His Asp Pro Thr Thr Arg Arg Gln Phe Ile
 245 250 255
- Val Ile Glu Thr Ser Glu Pro Phe Thr Lys Leu Thr Leu Thr Ala Ala 260 265 270
- His Leu Val Phe Val Gly Asn Ser Ser Ala Ala Ser Gly Ile Thr Ala 275 280 285
- Thr Phe Ala Ser Asn Val Lys Pro Gly Asp Thr Val Leu Val Trp Glu 290 295 300
- Asp Thr Cys Glu Ser Leu Lys Ser Val Thr Val Lys Arg Ile Tyr Thr 305 310 315 320
- Glu Glu His Glu Gly Ser Phe Ala Pro Val Thr Ala His Gly Thr Ile 325 330 335
- Ile Val Asp Gln Val Leu Ala Ser Cys Tyr Ala Val Ile Glu Asn His 340 345 350
- Lys Trp Ala His Trp Ala Phe Ala Pro Val Arg Leu Cys His Lys Leu 355 360 365
- Met Thr Trp Leu Phe Pro Ala Arg Glu Ser Asn Val Asn Phe Gln Glu 370 375 380
- Asp Gly Ile His Trp Tyr Ser Asn Met Leu Phe His Ile Gly Ser Trp 385 390 395 400

Leu Leu Asp Arg Asp Ser Phe His Pro Leu Gly Ile Leu His Leu Ser 405 410 415

<220 <221	.> 14 !> D! !> D: !> C!	416 NA roson	ohila (141)		lano	gaste	er									
<400			(141.	3)												
atg	gat	aac		_					•	_			agt Ser	_	_	48
_			_		_		_		_		_		cag Gln 30			96
			-	_					_	_		_	aga Arg	_		144
	_	_	-					_	_	_	_		agc Ser			192
		_							-				gtc Val			240
_	_		_	_						_		_	cat His		_	288
_		_		_	_	_		_	. –				aat Asn 110			336
		_		-	-								atc Ile			384
_	_				_	-							agg Arg	_		432
		_	_										atg Met			480

	_	_	_		aag Lys 165				_							_	528
					cgg Arg	-	_	_						_			576
				_	gag Glu					_	_		_	_	_	_	624
					cgc Arg												672
					gga Gly				_			-					720
					gtc Val 245												768
				_	ccg Pro	_	_							_	_		816
	_	_		_	gag Glu							_				_	864
					gcc Ala												912
,				_	atg Met											-	960
	_			_	gtg Val 325	_	_	-		_	-						1008
		_	_	_	ctc Leu	_						_			_		1056
		_			gta Val		_	_									1104
					ttg Leu												1152

a 1

Leu						Thr					Ser			gcc Ala		Cys	1200		
														ctg	_		1248		
Tyr	Al	la '	Val	Ile	Asn 405	Ser	Gln	Ser	Leu	Ala 410	His	Trp	Gly	Leu	Ala 415	Pro			
	_	-	_	_		_	_	_					_	aag Lys 430	_	_	1296		
		s												cag Gln			1344		
		s '												tac Tyr	_		1392		
						cac His 470		tga									1416		
<210 <211 <212 <213	l > { 2 > }	47 PR'	1 T	bhila	a mel	lanog	gaste	er										,	
<211 <212 <213 <400	1 > 6 2 > 3 3 > 3	47 PR' Dro	1 T osop						Trp	Ala 10	Ser	Ala	Ala	Ser	Val 15	Thr		•	
<211 <212 <213 <400 Met	1 >	PR'Dro	1 T osop Asn	His	Ser 5	Ser	Val	Pro	_	10				Ser Gln 30	15			•	
<211 <212 <213 <400 Met 1 Cys	1 > 6 2 > 1 3 > 1 0 > 1 As ₁	PR'Dro	1 T osop Asn Ser	His Leu 20	Ser 5 Gly	Ser	Val Gln	Pro Met	Pro 25	10 Gln	Phe	Gln	Phe	Gln	15 Phe	Gln			
<211 <212 <213 <400 Met 1 Cys	1 > 6 2 > 1 3 > 1 0 > 1 Asy	PR'Dro	1 T osor Asn Ser Ile 35	His Leu 20 Arg	Ser 5 Gly Ser	Ser Cys Glu	Val Gln Leu	Pro Met His 40	Pro 25 Leu	10 Gln Arg	Phe Lys	Gln Pro	Phe Ala 45	Gln 30	15 Phe Arg	Gln Thr			
<211 <212 <213 <400 Met 1 Cys Leu Gln	1 > 6 2 > 1 3 > 1 0 > 1 Asy Let Th	PR'Dro	1 T osop Asn Ser Ile 35 Met	His Leu 20 Arg	Ser 5 Gly Ser	Ser Cys Glu	Val Gln Leu Ala 55	Pro Met His 40	Pro 25 Leu Thr	10 Gln Arg	Phe Lys Arg	Gln Pro Cys 60	Phe Ala 45 Leu	Gln 30 Arg	15 Phe Arg	Gln Thr Leu			
<211 <212 <213 <400 Met	1> 6 2> 1 3> 1 0> 1 Asj Th 5	PR'Dro	1 T osop Asn Ser Ile 35 Met	His Leu 20 Arg Val	Ser 5 Gly Ser His	Ser Cys Glu Ile Leu 70	Val Gln Leu Ala 55 Leu	Pro Met His 40 His	Pro 25 Leu Thr	10 Gln Arg Gln Val	Phe Lys Arg Leu 75	Gln Pro Cys 60 Pro	Phe Ala 45 Leu Met	Gln 30 Arg Ser	15 Phe Arg Arg	Gln Thr Leu Ser 80			
<211 <212 <213 <400 Met	1> 1 2> 1 3> 1 0> 1 Asy Let Th 5 Se	PR'Dro	1 Tosor Asn Ser Ile 35 Met Leu	His Leu 20 Arg Val	Ser 5 Gly Ser His Ala Cys 85	Ser Cys Glu Ile Leu 70 Gly	Val Gln Leu Ala 55 Leu Pro	Pro Met His 40 His	Pro 25 Leu Thr	10 Gln Arg Gln Val Gly 90	Phe Lys Arg Leu 75 Leu	Gln Pro Cys 60 Pro	Phe Ala 45 Leu Met	Gln 30 Arg Ser Val	15 Phe Arg Phe Arg 95	Gln Thr Leu Ser 80 Ala			
<211 <212 <213 <400 Met	1> 1 2> 1 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	PR'Dro	1 T OSOF Asn Ser Ile 35 Met Leu	His Leu 20 Arg Val Ser Tyr 100	Ser 5 Gly Ser His Ala Cys 85 Pro	Ser Cys Glu Ile 70 Gly Leu	Val Gln Leu Ala 55 Leu Pro	Pro Met His 40 His Leu Gly	Pro 25 Leu Thr Ile Arg	10 Gln Arg Gln Val Gly 90 Gln	Phe Lys Arg Leu 75 Leu Thr	Gln Pro Cys 60 Pro Gly	Phe Ala 45 Leu Met Arg	Gln 30 Arg Ser Val His	Phe Arg Phe Arg State Arg	Gln Thr Leu Ser 80 Ala			

Leu Phe Arg Asp Glu Glu Gly Thr Gly Ala Asp Gly Leu Met Ser Lys
145 150 150

Arg Cys Lys Glu Lys Leu Asn Val Leu Ala Tyr Ser Val Met Asn Glu 165 170 175

Trp Pro Gly Ile Arg Leu Leu Val Thr Glu Ser Trp Asp Glu Asp Tyr 180 185 190

His His Gly Gln Glu Ser Leu His Tyr Glu Gly Arg Ala Val Thr Ile 195 200 205

Ala Thr Ser Asp Arg Asp Gln Ser Lys Tyr Gly Met Leu Ala Arg Leu 210 215 220

Ala Val Glu Ala Gly Phe Asp Trp Val Ser Tyr Val Ser Arg Arg His 225 230 235 240

Ile Tyr Cys Ser Val Lys Ser Asp Ser Ser Ile Ser Ser His Val His
245 250 255

Gly Cys Phe Thr Pro Glu Ser Thr Ala Leu Leu Glu Ser Gly Val Arg 260 265 270

Lys Pro Leu Gly Glu Leu Ser Ile Gly Asp Arg Val Leu Ser Met Thr 275 280 285

Ala Asn Gly Gln Ala Val Tyr Ser Glu Val Ile Leu Phe Met Asp Arg 290 295 300

Asn Leu Glu Gln Met Gln Asn Phe Val Gln Leu His Thr Asp Gly Gly 305 310 315 320

Ala Val Leu Thr Val Thr Pro Ala His Leu Val Ser Val Trp Gln Pro 325 330 335

Glu Ser Gln Lys Leu Thr Phe Val Phe Ala His Arg Ile Glu Glu Lys 340 345 350

Asn Gln Val Leu Val Arg Asp Val Glu Thr Gly Glu Leu Arg Pro Gln 355 360 365

Arg Val Val Lys Leu Gly Ser Val Arg Ser Lys Gly Val Val Ala Pro 370 375 380

Leu Thr Arg Glu Gly Thr Ile Val Val Asn Ser Val Ala Ala Ser Cys 385 390 395 400

Tyr Ala Val Ile Asn Ser Gln Ser Leu Ala His Trp Gly Leu Ala Pro 405 410 415

Met Arg Leu Leu Ser Thr Leu Glu Ala Trp Leu Pro Ala Lys Glu Gln
420 425 430

Leu His Ser Ser Pro Lys Val Val Ser Ser Ala Gln Gln Asn Gly

```
435 440 445

Ile His Trp Tyr Ala Asn Ala Leu Tyr Lys Val Lys Asp Tyr Val Leu
```

460

455

Pro Gln Ser Trp Arg His Asp 465 470

<210> 21

450

<211> 522

<212> DNA

<213> Homo sapiens

<400> 21

tgcggaccgg gcaggggtt cgggaagagg aggcaccca aaaagctgac ccctttagcc 60 tacaagcagt ttatcccaa tgtggccgag aagaccctag gcgccagcgg aaggtatgaa 120 gggaagatct ccagaaactc cgagcgattt aaggaactca cccccaatta caaccccgac 180 atcatattta aggatgaaga aaacaccgga gcggacaggc tgatgactca gaggtgtaag 240 gacaagttga acgctttggc catctcggtg atgaaccagt ggccaggagt gaaactgcgg 300 gtgaccgagg gctgggacga agatggccac cactcagagg agtctctgca ctacgagggc 360 cgcgcagtgg acatcaccac gtctgaccgc gaccgcagca agtacggcat gctggcccgc 420 ctggcggtgg aggccggctt cgactgggtg tactacgagt ccaaggcaca tatccactgc 480 tcggtgaaag cagagaactc ggtggcggcc aaatcgggag gc 522

<210> 22

<211> 525

<212> DNA

<213> Homo sapiens

<400> 22

tgegggeegg gtegggtggt gggeageege eggegaeege eacgeaaact egtgeegete 60 geetacaage agtteageec eaatgtgee gagaagaeee tgggegeeag eggaegetat 120 gaaggeaaga tegetegeag eteegagege tteaaggage teaeeecaa ttacaateea 180 gaeateatet teaaggaega ggagaacaca ggegeegaee geeteatgae eeagegetge 240 aaggaeege tgaaeteget ggetateteg gtgatgaaee agtggeeegg tgtgaagetg 240 eggetggegg tggaeateae eacateagae egegaeegea ataagtatgg aetgetggeg 420 egettggeag tggaggeegg etttgaetgg gtgtattaee agteaaagge eeacgtgeat 480 tgeteegtea agteegagea eteggeegea geeaagaegg gegge 525

<210> 23

<211> 175

<212> PRT

<213> Homo sapiens

<400> 23

Cys Gly Pro Gly Arg Val Val Gly Ser Arg Arg Arg Pro Pro Arg Lys
1 10 15

Leu Val Pro Leu Ala Tyr Lys Gln Phe Ser Pro Asn Val Pro Glu Lys 20 25 30

Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly Lys Ile Ala Arg Ser Ser

Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe 50 55 60

Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg Leu Met Thr Gln Arg Cys 65 70 75 80

Lys Asp Arg Leu Asn Ser Leu Ala Ile Ser Val Met Asn Gln Trp Pro 85 90 95

Gly Val Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His His
100 105 110

Ser Glu Glu Ser Leu His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr 115 120 125

Ser Asp Arg Asp Arg Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala Val 130 135 140

Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu Ser Lys Ala His Val His 145 150 155 160

Cys Ser Val Lys Ser Glu His Ser Ala Ala Ala Lys Thr Gly Gly
165 170 175

<210> 24

<211> 174

<212> PRT

<213> Homo sapiens

<400> 24

Cys Gly Pro Gly Arg Gly Phe Gly Lys Arg Arg His Pro Lys Lys Leu
1 5 10 15

Thr Pro Leu Ala Tyr Lys Gln Phe Ile Pro Asn Val Ala Glu Lys Thr 20 25 30

Leu Gly Ala Ser Gly Arg Tyr Glu Gly Lys Ile Ser Arg Asn Ser Glu
35 40 45

Arg Phe Lys Glu Leu Thr Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys 50 55 60

Asp Glu Glu Asn Thr Gly Ala Asp Arg Leu Met Thr Gln Arg Cys Lys 65 70 75 80

Asp Lys Leu Asn Ala Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly 85 90 95

Val Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His His Ser

Glu Glu Ser Leu His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser 115 120 125 Asp Arg Asp Arg Ser Lys Tyr Gly Met Leu Ala Arg Leu Ala Val Glu 130 135 140

Ala Gly Phe Asp Trp Val Tyr Tyr Glu Ser Lys Ala His Ile His Cys 145 150 155 160

Ser Val Lys Ala Glu Asn Ser Val Ala Ala Lys Ser Gly Gly
165 170

<210> 25

<211> 176

<212> PRT

<213> Homo sapiens

<400> 25

Cys Gly Pro Gly Arg Gly Pro Val Gly Arg Arg Arg Tyr Ala Arg Lys
1 5 10 15

Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe Val Pro Gly Val Pro Glu 20 25 30

Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu Gly Arg Val Ala Arg Gly
35 40 45

Ser Glu Arg Phe Arg Asp Leu Val Pro Asn Tyr Asn Pro Asp Ile Ile 50 55 60

Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp Arg Leu Met Thr Glu Arg
65 75 80

Cys Lys Glu Arg Val Asn Ala Leu Ala Ile Ala Val Met Asn Met Trp 85 90 95

Pro Gly Val Arg Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His
100 105 110

His Ala Gln Asp Ser Leu His Tyr Glu Gly Arg Ala Leu Asp Ile Thr 115 120 125

Thr Ser Asp Arg Asp Arg Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala 130 135 140

Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu Ser Arg Asn His Val 145 150 155 160

His Val Ser Val Lys Ala Asp Asn Ser Leu Ala Val Arg Ala Gly Gly
165 170 175

<210> 26

<211> 175

<212> PRT

<213> Artificial Sequence

```
<220>
<223> Description of Artificial Sequence: Consensus
      sequence
<220>
<221> SITE
<222> (1)
<223> Xaa=Cys that may be modified, altered or
      substituted within another moiety or series of
      moieties as described herein
<220>
<221> SITE
<222> (6)
<223> Xaa=Val or Gly
<220>
<221> SITE
<222> (7)
<223> Xaa=Val, Glu or Pro
<220>
<221> SITE
<222> (8)
<223> Xaa=Gly or Val
<220>
<221> SITE
<222> (9)
<223> Xaa=Ser or Gly
<220>
<221> SITE
<222> (10)
<223> Xaa=Arg or Lys
<220>
<221> SITE
<222> (13)
<223> Xaa=Pro, His or Tyr
<220>
<221> SITE
<222> (14)
<223> Xaa=Pro or Ala
<220>
<221> SITE
<222> (15)
<223> Xaa=Arg or Lys
<220>
<221> SITE
<222> (17)
```

<223> Xaa=any amino acid

```
<220>
<221> SITE
<222> (19)
<223> Xaa=Val or Thr
<220>
<221> SITE
<222> (22)
<223> Xaa=Ala or Leu
<220>
<221> SITE
<222> (27)
<223> Xaa=Ser, Ile or Val
<220>
<221> SITE
<222> (29)
<223> Xaa=Asn or Gly
<220>
<221> SITE
<222> (31)
<223> Xaa=Pro ar Ala
<220>
<221> SITE
<222> (41)
<223> Xaa=Tyr or Ala
<220>
<221> SITE
<222> (45)
<223> Xaa=Ile or Val
<220>
<221> SITE
<222> (46)
<223> Xaa=Ala or Ser
<220>
<221> SITE
<222> (48)
<223> Xaa=Ser, Asn or Gly
<220>
<221> SITE
<222> (54)
<223> Xaa=Glu or Asp
<220>
<221> SITE
<222> (56)
<223> Xaa=Thr or Val
<220>
<221> SITE
```

```
<222> (71)
<223> Xaa=Thr or Ser
<220>
<221> SITE
<222> (79)
<223> Xaa=Gln or Glu
<220>
<221> SITE
<222> (83)
<223> Xaa=Asp or Glu
<220>
<221> SITE
<222> (84)
<223> Xaa=Arg or Lys
<220>
<221> SITE
<222> (85)
<223> Xaa=Leu or Val
<220>
<221> SITE
<222> (91)
<223> Xaa=Ser or Ala
<220>
<221> SITE
<222> (95)
<223> Xaa=Gln or Met
<220>
<221> SITE
<222> (114)
<223> Xaa=Ser or Ala
<220>
<221> SITE
<222> (115)
<223> Xaa=Glu or Gln
<220>
<221> SITE
<222> (116)
<223> Xaa=Glu or Asp
<220>
<221> SITE
<222> (135)
<223> Xaa=Asn or Ser
<220>
<221> SITE
<222> (139)
<223> Xaa=Leu or Met
```

```
<220>
<221> SITE
<222> (157)
<223> Xaa=Lys or Arg
<220>
<221> SITE
<222> (158)
<223> Xaa=Ala or Asn
<220>
<221> SITE
<222> (160)
<223> Xaa=Val or Ile
<220>
<221> SITE
<222> (162)
<223> Xaa=Cys or Val
<220>
<221> SITE
<222> (166)
<223> Xaa=Ser or Ala
<220>
<221> SITE
<222> (167)
<223> Xaa=Glu or Asp
<220>
<221> SITE
<222> (168)
<223> Xaa=His or Asn
<220>
<221> SITE
<222> (169)
<223> Xaa=Ala, Val or Leu
<220>
<221> SITE
<222> (172)
<223> Xaa=Lys or Arg
<220>
<221> SITE
<222> (173)
<223> Xaa=Thr, Ser or Ala
<400> 26
Xaa Gly Pro Gly Arg Xaa Xaa Xaa Xaa Arg Arg Xaa Xaa Lys
 1
                  5
                                      10
                                                           15
Xaa Leu Xaa Pro Leu Xaa Tyr Lys Gln Phe Xaa Pro Xaa Val Xaa Glu
```

25

30

20

Lys Thr Leu Gly Ala Ser Gly Arg Xaa Glu Gly Lys Xaa Xaa Arg Xaa 35 45 Ser Glu Arg Phe Lys Xaa Leu Xaa Pro Asn Tyr Asn Pro Asp Ile Ile 50 55 60 Phe Lys Asp Glu Glu Asn Xaa Gly Ala Asp Arg Leu Met Thr Xaa Arg 65 70 75 Cys Lys Xaa Xaa Xaa Asn Ser Leu Ala Ile Xaa Val Met Asn Xaa Trp 95 85 90 Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His 100 105 110 His Xaa Xaa Xaa Ser Leu His Tyr Glu Gly Arg Ala Val Asp Ile Thr 115 125 120 Thr Ser Asp Arg Asp Arg Xaa Lys Tyr Gly Xaa Leu Ala Arg Leu Ala 130 135 140 Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu Ser Xaa Xaa His Xaa 145 150 155 160 His Xaa Ser Val Lys Xaa Xaa Xaa Xaa Ala Ala Xaa Xaa Gly Gly 170 175 165 <210> 27 <211> 528 <212> DNA <213> Homo sapiens <400> 27 tgcgggccgg gccggggcc ggttggccgg cgccgctatg cgcgcaagca gctcgtgccg 60 ctactctaca agcaatttgt gcccggcgtg ccagagcgga ccctgggcgc cagtgggcca 120 gcggagggga gggtggcaag gggctccgag cgcttccggg acctcgtgcc caactacaac 180 cccgacatca tcttcaagga tgaggagaac agtggagccg accgcctgat gaccgagcgt 240 tgtaaggagc gggtgaacgc tttggccatt gccgtgatga acatgtggcc cggagtgcgc 300 ctacgagtga ctgagggctg ggacgaggac ggccaccacg ctcaggattc actccactac 360 gaaggccgtg ctttggacat cactacgtct gaccgcgacc gcaacaagta tgggttgctg 420 gegegeeteg cagtggaage eggettegae tgggtetaet aegagteeeg caaccaegte 480 cacgtgtcgg tcaaagctga taactcactg gcggtccggg cgggcggc 528 <210> 28 <211> 684 <212> DNA <213> Homo sapiens <400> 28 gtcgacaaaa ctcacacatg cccaccgtgc ccagcacctg aactcctggg gggaccgtca 60 gtetteetet teececcaaa acceaaggae acceteatga teteceggae eeetgaggte 120 acatgcgtgg tggtggacgt gagccacgaa gaccctgagg tcaagttcaa ctggtacgtg 180

gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta ccagagcacg 240

```
taccgtgtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg caaggagtac 300
aagtgcaagg tctccaacaa agccctccca gcccccatcg agaaaaccat ctccaaagcc 360
aaagggcagc cccgagaacc acaggtgtac accctgcccc catcccggga tgagctgacc 420
aagaaccagg tcagcctgac ctgcctggtc aaaggcttct atcccagcga catcgccgtg 480
gagtgggaga gcaatgggca gccggagaac aactacaaga ccacgcctcc cgtgttggac 540
tecgaegget cettetteet etacageaag etcacegtgg acaagageag gtggeageag 600
gggaacgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta cacgcagaag 660
agectetece tgtetecegg gaaa
                                                                   684
<210> 29
<211> 687
<212> DNA
<213> Homo sapiens
<400> 29
gtcgacgtgc ccagggattg tggttgtaag ccttgcatat gtacagtccc agaagtatca 60
tetgtettea tettecece aaageeeaag gatgtgetea eeattaetet gaeteetaag 120
gtcacgtgtg ttgtggtaga catcagcaag gatgatcccg aggtccagtt cagctggttt 180
gtagatgatg tggaggtgca cacagctcag acgcaaccac gggaagagca gttccaaagc 240
actttccgct cagtcagtga acttcccatc atgcaccagg actggctcaa tggcaaggag 300
ttcaaatgca gggtcaacag tgcagctttc cctgccccca tcgagaaaac catctccaaa 360
accaaaggca gaccgaaggc tccacaggtg tacaccattc cacctcccaa ggagcagatg 420
gccaaggata aagtcagtct gacctgcatg ataacagact tcttccctga agacattact 480
gtggagtggc agtggaatgg gcagccagcg gagaactaca agaacactca gcccatcatg 540
gacacagatg gctcttactt cgtctacagc aagctcaatg tgcagaagag caactgggag 600
gcaggaaata ctttcacctg ctctgtgtta catgagggcc tgcacaacca ccatactgag 660
aagagcctct cccactctcc tggtaaa
                                                                  687
<210> 30
<211> 702
<212> DNA
<213> Homo sapiens
<400> 30
gtcgacccca gagggcccac aatcaagccc tgtcctccat gcaaatgccc agcacctaac 60
ctcttgggtg gaccatccgt cttcatcttc cctccaaaga tcaaggatgt actcatgatc 120
tccctgagcc ccatagtcac atgtgtggtg gtggatgtga gcgaggatga cccagatgtc 180
cagatcaget ggtttgtgaa caacgtggaa gtacacacag ctcagacaca aacccataga 240
gaggattacc aaagtacact tcgggtggtc agtgccctcc ccatccaqca ccaggactgg 300
atgagtggca aggagttcaa atgcaaggtc aacaacaaag acctcccagc gcccatcgag 360
agaaccatct caaaacccaa agggtcagta agagctccac aggtatatgt cttgcctcca 420
ccagaagaag agatgactaa gaaacaggtc actctgacct gcatggtgac agacttcatg 480
cctgaagaca tttacgtgga gtggaccaac aacgggaaaa cagagctaaa ctacaagaac 540
actgaaccag teetggacte tgatggttet taetteatgt acageaaget gagagtggaa 600
aagaagaact gggtggaaag aaatagctac tcctgttcag tggtccacga gggtctgcac 660
aatcaccaca cgactaagag cttctcccgg actccgggta aa
                                                                  702
<210> 31
<211> 9776
<212> DNA
<213> Plasmid P55
<400> 31
gatctaacat ccaaagacga aaggttgaat gaaacctttt tgccatccga catccacagg 60
```

```
tccattctca cacataagtg ccaaacgcaa caggagggga tacactagca gcagaccgtt 120
gcaaacgcag gacctccact cctcttctcc tcaacaccca cttttgccat cgaaaaacca 180
gcccagttat tgggcttgat tggagctcgc tcattccaat tccttctatt aggctactaa 240
caccatgact ttattagcct gtctatcctg gcccccctgg cgaggttcat gtttgtttat 300
ttccgaatgc aacaagctcc gcattacacc cgaacatcac tccagatgag ggctttctga 360
gtgtggggtc aaatagtttc atgttcccca aatggcccaa aactgacagt ttaaacgctg 420
tcttggaacc taatatgaca aaagcgtgat ctcatccaag atgaactaag tttggttcgt 480
tgaaatgcta acggccagtt ggtcaaaaag aaacttccaa aagtcgccat accgtttgtc 540
ttgtttggta ttgattgacg aatgctcaaa aataatctca ttaatgctta gcgcagtctc 600
tctatcgctt ctgaaccccg gtgcacctgt gccgaaacgc aaatggggaa acacccgctt 660
tttggatgat tatgcattgt ctccacattg tatgcttcca agattctggt gggaatactg 720
ctgatagcct aacgttcatg atcaaaattt aactgttcta acccctactt gacagcaata 780
tataaacaga aggaagctgc cctgtcttaa accttttttt ttatcatcat tattagctta 840
ctttcataat tgcgactggt tccaattgac aagcttttga ttttaacgac ttttaacgac 900
aacttgagaa gatcaaaaaa caactaatta ttcgaaggat ccaaacgatg agatttcctt 960
caatttttac tgcagtttta ttcgcagcat cctccgcatt agctgctcca gtcaacacta 1020
caacagaaga tgaaacggca caaattccgg ctgaagctgt catcggttac tcagatttag 1080
aaggggattt cgatgttgct gttttgccat tttccaacag cacaaataac gggttattgt 1140
ttataaatac tactattgcc agcattgctg ctaaagaaga aggggtatct ctcgagaaaa 1200
gatgcggacc gggcaggggg ttcgggaaga ggaggcaccc caaaaagctg acccctttag 1260
cctacaagca gtttatcccc aatgtggccg agaagaccct aggcgccagc ggaaggtatg 1320
aagggaagat ctccagaaac tccgagcgat ttaaggaact cacccccaat tacaaccccg 1380
acatcatatt taaggatgaa gaaaacaccg gagcggacag gctgatgact cagaggtgta 1440
aggacaagtt gaacgctttg gccatctcgg tgatgaacca gtggccagga gtgaaactgc 1500
gggtgaccga gggctgggac gaagatggcc accactcaga ggagtctctg cactacgagg 1560
gccgcgcagt ggacatcacc acgtctgacc gcgaccgcag caagtacggc atgctggccc 1620
gcctggcggt ggaggccggc ttcgactggg tgtactacga gtccaaggca catatccact 1680
gctcggtgaa agcagagaac tcggtggcgg ccaaatcggg aggctgattc gcggccgcga 1740
attaattcgc cttagacatg actgttcctc agttcaagtt gggcacttac gagaagaccg 1800
gtcttgctag attctaatca agaggatgtc agaatgccat ttgcctgaga gatgcaggct 1860
tcatttttga tacttttta tttgtaacct atatagtata ggattttttt tgtcattttg 1920
tttcttctcg tacgagettg ctcctgatca geetateteg cagetgatga atatettgtg 1980
gtaggggttt gggaaaatca ttcgagtttg atgtttttct tggtatttcc cactcctctt 2040
cagagtacag aagattaagt gagaagttcg tttgtgcaag cttatcgata agctttaatg 2100
cggtagttta tcacagttaa attgctaacg cagtcaggca ccgtgtatga aatctaacaa 2160
tgcgctcatc gtcatcctcg gcaccgtcac cctggatgct gtaggcatag gcttggttat 2220
gccggtactg ccgggcctct tgcgggatat cgtccattcc gacagcatcg ccagtcacta 2280
tggcgtgctg ctagcgctat atgcgttgat gcaatttcta tgcgcacccg ttctcggagc 2340
actgtccgac cgctttggcc gccgcccagt cctgctcgct tcgctacttg gagccactat 2400
cgactacgcg atcatggcga ccacacccgt cctgtggatc tatcgaatct aaatgtaagt 2460
taaaatetet aaataattaa ataagteeca gttteteeat acgaacetta acageattge 2520
ggtgagcatc tagaccttca acagcagcca gatccatcac tgcttggcca atatgtttca 2580
gtccctcagg agttacgtct tgtgaagtga tgaacttctg gaaggttgca gtgttaactc 2640
cgctgtattg acgggcatat ccgtacgttg gcaaagtgtg gttggtaccg gaggagtaat 2700
ctccacaact ctctggagag taggcaccaa caaacacaga tccagcgtgt tgtacttgat 2760
caacataaga agaagcattc tcgatttgca ggatcaagtg ttcaggagcg tactgattgg 2820
acatttccaa agcctgctcg taggttgcaa ccgatagggt tgtagagtgt gcaatacact 2880
tgcgtacaat ttcaaccett ggcaactgca cagettggtt gtgaacagca tettcaatte 2940
tggcaagctc cttgtctgtc atatcgacag ccaacagaat cacctgggaa tcaataccat 3000
gttcagcttg agcagaaggt ctgaggcaac gaaatctgga tcagcgtatt tatcagcaat 3060
aactagaact tcagaaggcc cagcaggcat gtcaatacta cacagggctg atgtgtcatt 3120
ttgaaccatc atcttggcag cagtaacgaa ctggtttcct ggaccaaata ttttgtcaca 3180
cttaggaaca gtttctgttc cgtaagccat agcagctact gcctgggcgc ctcctgctag 3240
cacgatacac ttagcaccaa ccttgtgggc aacgtagatg acttctgggg taagggtacc 3300
atccttctta ggtggagatg caaaaacaat ttctttgcaa ccagcaactt tggcaggaac 3360
acccagcatc agggaagtgg aaggcagaat tgcggttcca ccaggaatat agaggccaac 3420
tttctcaata ggtcttgcaa aacgagagca gactacacca gggcaagtct caacttgcaa 3480
```

```
cgtctccgtt agttgagctt catggaattt cctgacgtta tctatagaga gatcaatggc 3540
tctcttaacg ttatctggca attgcataag ttcctctggg aaaggagctt ctaacacagg 3600
tgtcttcaaa gcgactccat caaacttggc agttagttct aaaagggctt tgtcaccatt 3660
ttgacgaaca ttgtcgacaa ttggtttgac taattccata atctgttccg ttttctggat 3720
aggacgacga agggcatctt caatttcttg tgaggaggcc ttagaaacgt caattttgca 3780
caattcaata cgaccttcag aagggacttc tttaggtttg gattcttctt taggttgttc 3840
cttggtgtat cctggcttgg catctccttt ccttctagtg acctttaggg acttcatatc 3900
caggtttctc tccacctcgt ccaacgtcac accgtacttg gcacatctaa ctaatgcaaa 3960
ataaaataag tcagcacatt cccaggctat atcttccttg gatttagctt ctgcaagttc 4020
atcagcttcc tccctaattt tagcgttcaa acaaaacttc gtcgtcaaat aaccgtttgg 4080
tataagaacc ttctggagca ttgctcttac gatcccacaa ggtgcttcca tggctctaag 4140
accetttgat tggccaaaac aggaagtgeg ttccaagtga cagaaaccaa cacctgtttg 4200
ttcaaccaca aatttcaagc agtctccatc acaatccaat tcgataccca gcaacttttg 4260
agttcgtcca gatgtagcac ctttatacca caaaccgtga cgacgagatt ggtagactcc 4320
agtttgtgtc cttatagcct ccggaataga ctttttggac gagtacacca ggcccaacga 4380
gtaattagaa gagtcagcca ccaaagtagt gaatagacca tcggggcggt cagtagtcaa 4440
agacgccaac aaaatttcac tgacagggaa ctttttgaca tcttcagaaa gttcgtattc 4500
agtagtcaat tgccgagcat caataatggg gattatacca gaagcaacag tggaagtcac 4560
atctaccaac tttgcggtct cagaaaaagc ataaacagtt ctactaccgc cattagtgaa 4620
acttttcaaa tcgcccagtg gagaagaaaa aggcacagcg atactagcat tagcgggcaa 4680
ggatgcaact ttatcaacca gggtcctata gataacccta gcgcctggga tcatcctttg 4740
gacaactett tetgecaaat etaggteeaa aateaettea ttgataeeat tattgtaeaa 4800
cttgagcaag ttgtcgatca gctcctcaaa ttggtcctct gtaacggatg actcaacttg 4860
cacattaact tgaagctcag tcgattgagt gaacttgatc aggttgtgca gctggtcagc 4920
agcataggga aacacggctt ttcctaccaa actcaaggaa ttatcaaact ctgcaacact 4980
tgcgtatgca ggtagcaagg gaaatgtcat acttgaagtc ggacagtgag tgtagtcttg 5040
agaaattctg aagccgtatt tttattatca gtgagtcagt catcaggaga tcctctacgc 5100
cggacgcatc gtggccgacc tgcaggtcgg catcaccggc gccacaggtg cggttgctgg 5160
cgcctatatc gccgacatca ccgatgggga agatcgggct cgccacttcg ggctcatgag 5220
cgcttgtttc ggcgtgggta tggtggcagg ccccgtggcc gggggactgt tgggcgccat 5280
ctccttggac ctgcaggggg gggggggaa agccacgttg tgtctcaaaa tctctgatgt 5340
tacattgcac aagataaaaa tatatcatca tgaacaataa aactgtctgc ttacataaac 5400
agtaatacaa ggggtgttat gagccatatt caacgggaaa cgtcttgctc aaggccgcga 5460
ttaaattcca acatggatgc tgatttatat gggtataaat gggctcgcga taatgtcggg 5520
caatcaggtg cgacaatcta tcgattgtat gggaagcccg atgcgccaga gttgtttctg 5580
aaacatggca aaggtagcgt tgccaatgat gttacagatg agatggtcag actaaactgg 5640
ctgacggaat ttatgcctct tccgaccatc aagcatttta tccgtactcc tgatgatgca 5700
tggttactca ccactgcgat ccccgggaaa acagcattcc aggtattaga agaatatcct 5760
gattcaggtg aaaatattgt tgatgcgctg gcagtgttcc tgcgccggtt gcattcgatt 5820
cctgtttgta attgtccttt taacagcgat cgcgtatttc gtctcgctca ggcgcaatca 5880
cgaatgaata acggtttggt tgatgcgagt gattttgatg acgagcgtaa tggctggcct 5940
gttgaacaag tctggaaaga aatgcataag cttttgccat tctcaccgga ttcagtcgtc 6000
actcatggtg atttctcact tgataacctt atttttgacg aggggaaatt aataggttgt 6060
attgatgttg gacgagtcgg aatcgcagac cgataccagg atcttgccat cctatggaac 6120
tgcctcggtg agttttctcc ttcattacag aaacggcttt ttcaaaaata tggtattgat 6180
aatcctgata tgaataaatt gcagtttcat ttgatgctcg atgagttttt ctaatcagaa 6240
ttggttaatt ggttgtaaca ctggcagagc attacgctga cttgacggga cggcggcttt 6300
gttgaataaa tcgaactttt gctgagttga aggatcagat cacgcatctt cccgacaacg 6360
cagaccgttc cgtggcaaag caaaagttca aaatcaccaa ctggtccacc tacaacaaag 6420
ctctcatcaa ccgtggctcc ctcactttct ggctggatga tggggcgatt caggcctggt 6480
atgagtcagc aacacettet teacgaggea gacetcageg ecceecece ectgeaggte 6540
ccacggcggc ggtgctcaac ggcctcaacc tactactggg ctgcttccta atgcaggagt 6600
cgcataaggg agagcgtcga gtatctatga ttggaagtat gggaatggtg atacccgcat 6660
tcttcagtgt cttgaggtct cctatcagat tatgcccaac taaagcaacc ggaggaggag 6720
atttcatggt aaatttctct gacttttggt catcagtaga ctcgaactgt gagactatct 6780
cggttatgac agcagaaatg tccttcttgg agacagtaaa tgaagtccca ccaataaaga 6840
aatccttgtt atcaggaaca aacttcttgt ttcgaacttt ttcggtgcct tgaactataa 6900
```

```
aatgtagagt ggatatgtcg ggtaggaatg gagcgggcaa atgcttacct tctggacctt 6960
caagaggtat gtagggtttg tagatactga tgccaacttc agtgacaacg ttgctatttc 7020
gttcaaacca ttccgaatcc agagaaatca aagttgtttg tctactattg atccaagcca 7080
gtgcggtctt gaaactgaca atagtgtgct cgtgttttga ggtcatcttt gtatgaataa 7140
atctagtctt tgatctaaat aatcttgacg agccaaggcg ataaataccc aaatctaaaa 7200
ctcttttaaa acgttaaaag gacaagtatg tctgcctgta ttaaacccca aatcagctcg 7260
tagtctgatc ctcatcaact tgaggggcac tatcttgttt tagagaaatt tgcggagatg 7320
cgatatcgag aaaaaggtac gctgatttta aacgtgaaat ttatctcaag atctctgcct 7380
cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg cagctcccgg agacggtcac 7440
agettgtetg taageggatg eegggageag acaageeegt eagggegegt eagegggtgt 7500
tggcgggtgt cggggcgcag ccatgaccca gtcacgtagc gatagcggag tgtatactgg 7560
cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatgcg gtgtgaaata 7620
ccgcacagat gcgtaaggag aaaataccgc atcaggcgct cttccgcttc ctcgctcact 7680
gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat cagctcactc aaaggcggta 7740
atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag 7800
caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc 7860
cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta 7920
taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg 7980
ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct ttctcaatgc 8040
tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac 8100
gaaccccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct tgagtccaac 8160
ccggtaagac acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg 8220
aggtatgtag gcggtgctac agagttcttg aagtggtggc ctaactacgg ctacactaga 8280
aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt 8340
agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gtttttttgt ttgcaagcag 8400
cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct 8460
gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 8520
atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat 8580
gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 8640
tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatacgg 8700
gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg ctcaccggct 8760
ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggtcctgca 8820
actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg 8880
ccagttaata gtttgcgcaa cgttgttgcc attgctgcag gcatcgtggt gtcacgctcg 8940
tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc 9000
cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt cagaagtaag 9060
ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg 9120
ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt ctgagaatag 9180
tgtatgcggc gaccgagttg ctcttgcccg gcgtcaacac gggataatac cgcgccacat 9240
agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa actctcaagg 9300
atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa ctgatcttca 9360
gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca 9420
aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct ttttcaatat 9480
tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag 9540
aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa 9600
gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag gccctttcgt 9660
cttcaagaat taattctcat gtttgacagc ttatcatcga taagctgact catgttggta 9720
ttgtgaaata gacgcagatc gggaacactg aaaaataaca gttattattc gagatc
                                                                  9776
```

```
<210> 32
<211> 10491
<212> DNA
<213> Plasmid pUB114
```

<400> 32

gatctaacat ccaaagacga aaggttgaat gaaacctttt tgccatccga catccacagg 60

```
tccattctca cacataagtg ccaaacgcaa caggagggga tacactagca gcagaccgtt 120
gcaaacgcag gacctccact cctcttctcc tcaacaccca cttttgccat cgaaaaacca 180
gcccagttat tgggcttgat tggagctcgc tcattccaat tccttctatt aggctactaa 240
caccatgact ttattagcct gtctatcctg gccccctgg cgaggttcat gtttgtttat 300
ttccgaatgc aacaagctcc gcattacacc cgaacatcac tccagatgag ggctttctga 360
gtgtggggtc aaatagtttc atgttcccca aatggcccaa aactgacagt ttaaacgctg 420
tcttggaacc taatatgaca aaagcgtgat ctcatccaag atgaactaag tttggttcgt 480
tgaaatgcta acggccagtt ggtcaaaaag aaacttccaa aagtcgccat accgtttgtc 540
ttgtttggta ttgattgacg aatgctcaaa aataatctca ttaatgctta gcgcagtctc 600
tctatcgctt ctgaaccccg gtgcacctgt gccgaaacgc aaatggggaa acacccgctt 660
tttggatgat tatgcattgt ctccacattg tatgcttcca agattctggt gggaatactg 720
ctgatagcct aacgttcatg atcaaaattt aactgttcta acccctactt gacagcaata 780
tataaacaga aggaagctgc cctgtcttaa acctttttt tatcatcatt attagcttac 840
tttcataatt gcgactggtt ccaattgaca agcttttgat tttaacgact tttaacgaca 900
acttgagaag atcaaaaaac aactaattat tcgaaggatc caaacgatga gatttccttc 960
aatttttact gcagttttat tcgcagcatc ctccgcatta gctgctccag tcaacactac 1020
aacagaagat gaaacggcac aaattccggc tgaagctgtc atcggttact cagatttaga 1080
aggggatttc gatgttgctg ttttgccatt ttccaacagc acaaataacg ggttattgtt 1140
tataaatact actattgcca gcattgctgc taaagaagaa ggggtatctc tcgagaaaag 1200
atgcggaccg ggcaggggt tcgggaagag gaggcacccc aaaaagctga cccctttagc 1260
ctacaagcag tttatcccca atgtggccga gaagacccta ggcgccagcg gaaggtatga 1320
agggaagatc tccagaaact ccgagcgatt taaggaactc acccccaatt acaaccccga 1380
catcatattt aaggatgaag aaaacaccgg agcggacagg ctgatgactc agaggtgtaa 1440
ggacaagttg aacgetttgg ccatcteggt gatgaaccag tggccaggag tgaaactgcg 1500
ggtgaccgag ggctgggacg aagatggcca ccactcagag gagtctctgc actacgaggg 1560
ccgcgcagtg gacatcacca cgtctgaccg cgaccgcagc aagtacggca tgctggcccg 1620
cctggcggtg gaggccggct tcgactgggt gtactacgag tccaaggcac atatccactg 1680
ctcggtgaaa gcagagaact cggtggcggc caaatcggga ggcgtcgacg tgcccaggga 1740
ttgtggttgt aagcettgea tatgtacagt eecagaagta teatetgtet teatetteee 1800
cccaaagccc aaggatgtgc tcaccattac tctgactcct aaggtcacgt gtgttgtggt 1860
agacatcage aaggatgate eegaggteea gtteagetgg tttgtagatg atgtggaggt 1920
gcacacaget cagacgcaac cacgggaaga gcagttecaa agcaetttee geteagteag 1980
tgaacttccc atcatgcacc aggactggct caatggcaag gagttcaaat gcagggtcaa 2040
cagtgcagct ttccctgccc ccatcgagaa aaccatctcc aaaaccaaag gcagaccgaa 2100
ggctccacag gtgtacacca ttccacctcc caaggagcag atggccaagg ataaagtcag 2160
tctgacctgc atgataacag acttcttccc tgaagacatt actgtggagt ggcagtggaa 2220
tgggcagcca gcggagaact acaagaacac tcagcccatc atggacacag atggctctta 2280
cttcgtctac agcaagctca atgtgcagaa gagcaactgg gaggcaggaa atactttcac 2340
ctgctctgtg ttacatgagg gcctgcacaa ccaccatact gagaagagcc tctcccactc 2400
tcctggtaaa tgatcccagt gtccttggag ccctctggtc ctacagcggc cgcgaattaa 2460
ttcgccttag acatgactgt tcctcagttc aagttgggca cttacgagaa gaccggtctt 2520
gctagattct aatcaagagg atgtcagaat gccatttgcc tgagagatgc aggcttcatt 2580
tttgatactt ttttatttgt aacctatata gtataggatt ttttttgtca ttttgtttct 2640
tetegtacga gettgeteet gateageeta tetegeaget gatgaatate ttgtggtagg 2700
ggtttgggaa aatcattcga gtttgatgtt tttcttggta tttcccactc ctcttcagag 2760
tacagaagat taagtgagaa gttcgtttgt gcaagcttat cgataagctt taatgcggta 2820
gtttatcaca gttaaattgc taacgcagtc aggcaccgtg tatgaaatct aacaatgcgc 2880
tcatcgtcat cctcggcacc gtcaccctgg atgctgtagg cataggcttg gttatgccgg 2940
tactgccggg cctcttgcgg gatatcgtcc attccgacag catcgccagt cactatggcg 3000
tgctgctagc gctatatgcg ttgatgcaat ttctatgcgc acccgttctc ggagcactgt 3060
acgcgatcat ggcgaccaca cccgtcctgt ggatctatcg aatctaaatg taagttaaaa 3180
tetetaaata attaaataag teecagttte teeataegaa eettaacage attgeggtga 3240
gcatctagac cttcaacagc agccagatcc atcactgctt ggccaatatg tttcagtccc 3300
tcaggagtta cgtcttgtga agtgatgaac ttctggaagg ttgcagtgtt aactccgctg 3360
tattgacggg catatccgta cgttggcaaa gtgtggttgg taccggagga gtaatctcca 3420
caactetetg gagagtagge accaacaaac acagatecag egtgttgtae ttgateaaca 3480
```

taagaagaag cattctcgat ttgcaggatc aagtgttcag gagcgtactg attggacatt 3540 tccaaagcct gctcgtaggt tgcaaccgat agggttgtag agtgtgcaat acacttgcgt 3600 acaatttcaa cccttggcaa ctgcacagct tggttgtgaa cagcatcttc aattctggca 3660 ageteettgt etgteatate gacageeaac agaateacet gggaateaat accatgttea 3720 gcttgagcag aaggtctgag gcaacgaaat ctggatcagc gtatttatca gcaataacta 3780 gaacttcaga aggcccagca ggcatgtcaa tactacacag ggctgatgtg tcattttgaa 3840 ccatcatctt ggcagcagta acgaactggt ttcctggacc aaatattttg tcacacttag 3900 gaacagtttc tgttccgtaa gccatagcag ctactgcctg ggcgcctcct gctagcacga 3960 tacacttagc accaaccttg tgggcaacgt agatgacttc tggggtaagg gtaccatcct 4020 tcttaggtgg agatgcaaaa acaatttctt tgcaaccagc aactttggca ggaacaccca 4080 gcatcaggga agtggaaggc agaattgcgg ttccaccagg aatatagagg ccaactttct 4140 caataggtct tgcaaaacga gagcagacta caccagggca agtctcaact tgcaacgtct 4200 ccgttagttg agcttcatgg aatttcctga cgttatctat agagagatca atggctctct 4260 taacgttatc tggcaattgc ataagttcct ctgggaaagg agcttctaac acaggtgtct 4320 tcaaagcgac tccatcaaac ttggcagtta gttctaaaag ggctttgtca ccattttgac 4380 gaacattgtc gacaattggt ttgactaatt ccataatctg ttccgttttc tggataggac 4440 gacgaagggc atcttcaatt tcttgtgagg aggccttaga aacgtcaatt ttgcacaatt 4500 caatacgacc ttcagaaggg acttctttag gtttggattc ttctttaggt tgttccttgg 4560 tgtatcctgg cttggcatct cctttccttc tagtgacctt tagggacttc atatccaggt 4620 ttctctccac ctcgtccaac gtcacaccgt acttggcaca tctaactaat gcaaaataaa 4680 ataagtcagc acattcccag gctatatctt ccttggattt agcttctgca agttcatcag 4740 cttcctccct aattttagcg ttcaaacaaa acttcgtcgt caaataaccg tttggtataa 4800 gaacettetg gageattget ettacgatee cacaaggtge ttecatgget etaagaceet 4860 ttgattggcc aaaacaggaa gtgcgttcca agtgacagaa accaacacct gtttgttcaa 4920 ccacaaattt caagcagtct ccatcacaat ccaattcgat acccagcaac ttttgagttc 4980 gtccagatgt agcaccttta taccacaaac cgtgacgacg agattggtag actccagttt 5040 gtgtccttat agcctccgga atagactttt tggacgagta caccaggccc aacgagtaat 5100 tagaagagtc agccaccaaa gtagtgaata gaccatcggg gcggtcagta gtcaaagacg 5160 ccaacaaaat ttcactgaca gggaactttt tgacatcttc agaaagttcg tattcagtag 5220 tcaattgccg agcatcaata atggggatta taccagaagc aacagtggaa gtcacatcta 5280 ccaactttgc ggtctcagaa aaagcataaa cagttctact accgccatta gtgaaacttt 5340 tcaaatcgcc cagtggagaa gaaaaaggca cagcgatact agcattagcg ggcaaggatg 5400 caactttatc aaccagggtc ctatagataa ccctagcgcc tgggatcatc ctttggacaa 5460 ctctttctgc caaatctagg tccaaaatca cttcattgat accattattg tacaacttga 5520 gcaagttgtc gatcagctcc tcaaattggt cctctgtaac ggatgactca acttgcacat 5580 taacttgaag ctcagtcgat tgagtgaact tgatcaggtt gtgcagctgg tcagcagcat 5640 agggaaacac ggcttttcct accaaactca aggaattatc aaactctgca acacttgcgt 5700 atgcaggtag caagggaaat gtcatacttg aagtcggaca gtgagtgtag tcttgagaaa 5760 ttctgaagcc gtatttttat tatcagtgag tcagtcatca ggagatcctc tacgccggac 5820 gcatcgtggc cgacctgcag gtcggcatca ccggcgccac aggtgcggtt gctggcgcct 5880 atatcgccga catcaccgat ggggaagatc gggctcgcca cttcgggctc atgagcgctt 5940 gtttcggcgt gggtatggtg gcaggccccg tggccggggg actgttgggc gccatctcct 6000 tggacctgca ggggggggg gggaaagcca cgttgtgtct caaaatctct gatgttacat 6060 tgcacaagat aaaaatatat catcatgaac aataaaactg tctgcttaca taaacagtaa 6120 tacaaggggt gttatgagcc atattcaacg ggaaacgtct tgctcaaggc cgcgattaaa 6180 ttccaacatg gatgctgatt tatatgggta taaatgggct cgcgataatg tcgggcaatc 6240 aggtgcgaca atctatcgat tgtatgggaa gcccgatgcg ccagagttgt ttctgaaaca 6300 ggaatttatg cctcttccga ccatcaagca ttttatccgt actcctgatg atgcatggtt 6420 actcaccact gcgatccccg ggaaaacagc attccaggta ttagaagaat atcctgattc 6480 aggtgaaaat attgttgatg cgctggcagt gttcctgcgc cggttgcatt cgattcctgt 6540 ttgtaattgt ccttttaaca gcgatcgcgt atttcgtctc gctcaggcgc aatcacgaat 6600 gaataacggt ttggttgatg cgagtgattt tgatgacgag cgtaatggct ggcctgttga 6660 acaagtctgg aaagaaatgc ataagctttt gccattctca ccggattcag tcgtcactca 6720 tggtgatttc tcacttgata accttatttt tgacgagggg aaattaatag gttgtattga 6780 tgttggacga gtcggaatcg cagaccgata ccaggatctt gccatcctat ggaactgcct 6840 cqqtqaqttt tctccttcat tacagaaacg gctttttcaa aaatatggta ttgataatcc 6900

tgatatgaat aaattgcagt ttcatttgat gctcgatgag tttttctaat cagaattggt 6960 taattggttg taacactggc agagcattac gctgacttga cgggacggcg gctttgttga 7020 ataaatcgaa cttttgctga gttgaaggat cagatcacgc atcttcccga caacgcagac 7080 cgttccgtgg caaagcaaaa gttcaaaatc accaactggt ccacctacaa caaagctctc 7140 atcaaccgtg gctccctcac tttctggctg gatgatgggg cgattcaggc ctggtatgag 7200 tcagcaacac cttcttcacg aggcagacct cagcgccccc cccccctgc aggtcccacg 7260 geggeggtge teaacggeet caacetacta etgggetget teetaatgea ggagtegeat 7320 aagggagage gtegagtate tatgattgga agtatgggaa tggtgatace egeattette 7380 agtgtcttga ggtctcctat cagattatgc ccaactaaag caaccggagg aggagatttc 7440 atggtaaatt tetetgaett ttggteatea gtagaetega aetgtgagae tateteggtt 7500 atgacagcag aaatgtcctt cttggagaca gtaaatgaag tcccaccaat aaagaaatcc 7560 ttgttatcag gaacaaactt cttgtttcga actttttcgg tgccttgaac tataaaatgt 7620 agagtggata tgtcgggtag gaatggagcg ggcaaatgct taccttctgg accttcaaga 7680 ggtatgtagg gtttgtagat actgatgcca acttcagtga caacgttgct atttcgttca 7740 aaccattccg aatccagaga aatcaaagtt gtttgtctac tattgatcca agccagtgcg 7800 gtcttgaaac tgacaatagt gtgctcgtgt tttgaggtca tctttgtatg aataaatcta 7860 gtctttgatc taaataatct tgacgagcca aggcgataaa tacccaaatc taaaactctt 7920 ttaaaacgtt aaaaggacaa gtatgtctgc ctgtattaaa ccccaaatca gctcgtagtc 7980 tgatcctcat caacttgagg ggcactatct tgttttagag aaatttgcgg agatgcgata 8040 tcgagaaaaa ggtacgctga ttttaaacgt gaaatttatc tcaagatctc tgcctcgcgc 8100 gtttcggtga tgacggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt 8160 gtctgtaagc ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg 8220 ggtgtcgggg cgcagccatg acccagtcac gtagcgatag cggagtgtat actggcttaa 8280 ctatgcggca tcagagcaga ttgtactgag agtgcaccat atgcggtgtg aaataccgca 8340 cagatgcgta aggagaaaat accgcatcag gcgctcttcc gcttcctcgc tcactgactc 8400 gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 8460 gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaa gccagcaaaa 8520 ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 8580 cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 8640 ataccaggeg tttececetg gaageteest egtgegetet eetgtteega eeetgeeget 8700 taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc aatgctcacg 8760 ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 8820 ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 8880 aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 8940 tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac 9000 agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 9060 ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 9120 tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 9180 tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 9240 cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 9300 aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 9360 atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg 9420 cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga 9480 tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt 9540 atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt 9600 taatagtttg cgcaacgttg ttgccattgc tgcaggcatc gtggtgtcac gctcgtcgtt 9660 tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat 9720 gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc 9780 cgcagtgtta tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc 9840 cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat 9900 geggegaeeg agttgetett geeeggegte aacaegggat aataeegege cacatageag 9960 aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt 10020 acceptette agatecagtt egatetaace cactegtea eccaacteat etteageate 10080 ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa 10140 gggaataagg gcgacacgga aatgttgaat actcatactc ttcctttttc aatattattg 10200 aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 10260 taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac 10320 cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtcttca 10380 agaattaatt ctcatgtttg acagcttatc atcgataagc tgactcatgt tggtattgtg 10440 aaatagacgc agatcgggaa cactgaaaaa taacagttat tattcgagat c 10491

<210> 33 <211> 10512 <212> DNA <213> Plasmid pUB115

<400> 33

gatctaacat ccaaagacga aaggttgaat gaaacctttt tgccatccga catccacagg 60 tccattctca cacataagtg ccaaacgcaa caggagggga tacactagca gcagaccgtt 120 gcaaacgcag gacctccact cctcttctcc tcaacaccca cttttgccat cgaaaaacca 180 gcccagttat tgggcttgat tggagctcgc tcattccaat tccttctatt aggctactaa 240 caccatgact ttattagcct gtctatcctg gccccctgg cgaggttcat gtttgtttat 300 ttccgaatgc aacaagctcc gcattacacc cgaacatcac tccagatgag ggctttctga 360 gtgtggggtc aaatagtttc atgttcccca aatggcccaa aactgacagt ttaaacgctg 420 tcttggaacc taatatgaca aaagcgtgat ctcatccaag atgaactaag tttggttcgt 480 tgaaatgcta acggccagtt ggtcaaaaag aaacttccaa aagtcgccat accgtttgtc 540 ttgtttggta ttgattgacg aatgctcaaa aataatctca ttaatgctta gcgcagtctc 600 tctatcgctt ctgaaccccg gtgcacctgt gccgaaacgc aaatggggaa acacccgctt 660 tttggatgat tatgcattgt ctccacattg tatgcttcca agattctggt gggaatactg 720 ctgatagcct aacgttcatg atcaaaattt aactgttcta acccctactt gacagcaata 780 tataaacaga aggaagctgc cctgtcttaa accttttttt ttatcatcat tattagctta 840 ctttcataat tgcgactggt tccaattgac aagcttttga ttttaacgac ttttaacgac 900 aacttgagaa gatcaaaaaa caactaatta ttcgaaggat ccaaacgatg agatttcctt 960 caatttttac tgcagtttta ttcgcagcat cctccgcatt agctgctcca gtcaacacta 1020 caacagaaga tgaaacggca caaattccgg ctgaagctgt catcggttac tcagatttag 1080 aaggggattt cgatgttgct gttttgccat tttccaacag cacaaataac gggttattgt 1140 ttataaatac tactattgcc agcattgctg ctaaagaaga aggggtatct ctcgagaaaa 1200 gatgcggacc gggcaggggg ttcgggaaga ggaggcaccc caaaaagctg acccctttag 1260 cctacaagca gtttatcccc aatgtggccg agaagaccct aggcgccagc ggaaggtatg 1320 aagggaagat ctccagaaac tccgagcgat ttaaggaact cacccccaat tacaaccccg 1380 acatcatatt taaggatgaa gaaaacaccg gagcggacag gctgatgact cagaggtgta 1440 aggacaagtt gaacgctttg gccatctcgg tgatgaacca gtggccagga gtgaaactgc 1500 gggtgaccga gggctgggac gaagatggcc accactcaga ggagtctctg cactacgagg 1560 geogegeagt ggacateace aegtetgace gegacegeag caagtaegge atgetggeee 1620 gcctggcggt ggaggccggc ttcgactggg tgtactacga gtccaaggca catatccact 1680 gctcggtgaa agcagagaac tcggtggcgg ccaaatcggg aggcgtcgac cccagagggc 1740 ccacaatcaa gccctgtcct ccatgcaaat gcccagcacc taacctcttg ggtggaccat 1800 ccgtcttcat cttccctcca aagatcaagg atgtactcat gatctccctg agccccatag 1860 tcacatgtgt ggtggtggat gtgagcgagg atgacccaga tgtccagatc agctggtttg 1920 tgaacaacgt ggaagtacac acagctcaga cacaaaccca tagagaggat taccaaagta 1980 cacttegggt ggteagtgee etececatee ageaceagga etggatgagt ggeaaggagt 2040 tcaaatgcaa ggtcaacaac aaagacctcc cagcgcccat cgagagaacc atctcaaaac 2100 ccaaagggtc agtaagagct ccacaggtat atgtcttgcc tccaccagaa gaagagatga 2160 ctaagaaaca ggtcactctg acctgcatgg tgacagactt catgcctgaa gacatttacg 2220 tggagtggac caacaacggg aaaacagagc taaactacaa gaacactgaa ccagtcctgg 2280 actctgatgg ttcttacttc atgtacagca agctgagagt ggaaaagaag aactgggtgg 2340 aaagaaatag ctactcctgt tcagtggtcc acgagggtct gcacaatcac cacacgacta 2400 agagettete eeggaeteeg ggtaaatgag eteagatega tteeatggat eeteacatee 2460 caatccgcgg ccgcgaatta attcgcctta gacatgactg ttcctcagtt caagttgggc 2520 acttacgaga agaccggtct tgctagattc taatcaagag gatgtcagaa tgccatttgc 2580 ctgagagatg caggetteat ttttgataet tttttatttg taacetatat agtataggat 2640 tttttttgtc attttgtttc ttctcgtacg agcttgctcc tgatcagcct atctcgcagc 2700 tgatgaatat cttgtggtag gggtttggga aaatcattcg agtttgatgt ttttcttggt 2760

```
atttcccact cctcttcaga gtacagaaga ttaagtgaga agttcgtttg tgcaagctta 2820
tcgataagct ttaatgcggt agtttatcac agttaaattg ctaacgcagt caggcaccgt 2880
gtatgaaatc taacaatgcg ctcatcgtca tcctcggcac cgtcaccctg gatgctgtag 2940
gcataggett ggttatgeeg gtaetgeegg geetettgeg ggatategte catteegaea 3000
gcatcgccag tcactatggc gtgctgctag cgctatatgc gttgatgcaa tttctatgcg 3060
caccegttet eggageactg teegaceget ttggeegeeg ceeagteetg etegettege 3120
tacttggagc cactatcgac tacgcgatca tggcgaccac acccgtcctg tggatctatc 3180
gaatctaaat gtaagttaaa atctctaaat aattaaataa gtcccagttt ctccatacga 3240
accttaacag cattgcggtg agcatctaga ccttcaacag cagccagatc catcactgct 3300
tggccaatat gtttcagtcc ctcaggagtt acgtcttgtg aagtgatgaa cttctggaag 3360
gttgcagtgt taactccgct gtattgacgg gcatatccgt acgttggcaa agtgtggttg 3420
gtaccggagg agtaatctcc acaactctct ggagagtagg caccaacaaa cacagatcca 3480
gcgtgttgta cttgatcaac ataagaagaa gcattctcga tttgcaggat caagtgttca 3540
ggagcgtact gattggacat ttccaaagcc tgctcgtagg ttgcaaccga tagggttgta 3600
gagtgtgcaa tacacttgcg tacaatttca accettggca actgcacage ttggttgtga 3660
acagcatett caattetgge aageteettg tetgteatat egacageeaa cagaateace 3720
tgggaatcaa taccatgttc agcttgagca gaaggtctga ggcaacgaaa tctggatcag 3780
cgtatttatc agcaataact agaacttcag aaggcccagc aggcatgtca atactacaca 3840
gggctgatgt gtcattttga accatcatct tggcagcagt aacgaactgg tttcctggac 3900
caaatatttt gtcacactta ggaacagttt ctgttccgta agccatagca gctactgcct 3960
gggcgcctcc tgctagcacg atacacttag caccaacctt gtgggcaacg tagatgactt 4020
ctggggtaag ggtaccatcc ttcttaggtg gagatgcaaa aacaatttct ttgcaaccag 4080
caactttggc aggaacaccc agcatcaggg aagtggaagg cagaattgcg gttccaccag 4140
gaatatagag gccaactttc tcaataggtc ttgcaaaacg agagcagact acaccagggc 4200
aagteteaae ttgeaaegte teegttagtt gagetteatg gaattteetg aegttateta 4260
tagagagatc aatggctctc ttaacgttat ctggcaattg cataagttcc tctgggaaag 4320
gagcttctaa cacaggtgtc ttcaaagcga ctccatcaaa cttggcagtt agttctaaaa 4380
gggctttgtc accattttga cgaacattgt cgacaattgg tttgactaat tccataatct 4440
gttccgtttt ctggatagga cgacgaaggg catcttcaat ttcttgtgag gaggccttag 4500
aaacgtcaat tttgcacaat tcaatacgac cttcagaagg gacttcttta ggtttggatt 4560
ttagggactt catatccagg tttctctcca cctcgtccaa cgtcacaccg tacttggcac 4680
atctaactaa tgcaaaataa aataagtcag cacattccca ggctatatct tccttggatt 4740
tagcttctgc aagttcatca gcttcctccc taattttagc gttcaaacaa aacttcgtcg 4800
tcaaataacc gtttggtata agaaccttct ggagcattgc tcttacgatc ccacaaggtg 4860
cttccatggc tctaagaccc tttgattggc caaaacagga agtgcgttcc aagtgacaga 4920
aaccaacacc tgtttgttca accacaatt tcaagcagtc tccatcacaa tccaattcga 4980
tacccagcaa cttttgagtt cgtccagatg tagcaccttt ataccacaaa ccgtgacgac 5040
gagattggta gactccagtt tgtgtcctta tagcctccgg aatagacttt ttggacgagt 5100
acaccaggee caacgagtaa ttagaagagt cagecaccaa agtagtgaat agaccategg 5160
ggcggtcagt agtcaaagac gccaacaaaa tttcactgac agggaacttt ttgacatctt 5220
cagaaagttc gtattcagta gtcaattgcc gagcatcaat aatggggatt ataccagaag 5280
caacagtgga agtcacatct accaactttg cggtctcaga aaaagcataa acagttctac 5340
taccgccatt agtgaaactt ttcaaatcgc ccagtggaga agaaaaaggc acagcgatac 5400
tagcattagc gggcaaggat gcaactttat caaccagggt cctatagata accctagcgc 5460
ctgggatcat cctttggaca actctttctg ccaaatctag gtccaaaatc acttcattga 5520
taccattatt gtacaacttg agcaagttgt cgatcagctc ctcaaattgg tcctctgtaa 5580
cggatgactc aacttgcaca ttaacttgaa gctcagtcga ttgagtgaac ttgatcaggt 5640
tgtgcagctg gtcagcagca tagggaaaca cggcttttcc taccaaactc aaggaattat 5700
caaactctgc aacacttgcg tatgcaggta gcaagggaaa tgtcatactt qaagtcggac 5760
agtgagtgta gtcttgagaa attctgaagc cgtattttta ttatcagtga gtcagtcatc 5820
aggagateet etaegeegga egeategtgg eegaeetgea ggteggeate aeeggegeea 5880
caggtgcggt tgctggcgcc tatatcgccg acatcaccga tggggaagat cgggctcgcc 5940
acttcgggct catgagcgct tgtttcggcg tgggtatggt ggcaggcccc gtggccgggg 6000
gactgttggg cgccatctcc ttggacctgc aggggggggg ggggaaagcc acgttgtgtc 6060
tcaaaatctc tgatgttaca ttgcacaaga taaaaatata tcatcatgaa caataaaact 6120
gtctgcttac ataaacagta atacaagggg tgttatgagc catattcaac gggaaacgtc 6180
```

ttgctcaagg ccgcgattaa attccaacat ggatgctgat ttatatgggt ataaatgggc 6240 tcgcgataat gtcgggcaat caggtgcgac aatctatcga ttgtatggga agcccgatgc 6300 gccagagttg tttctgaaac atggcaaagg tagcgttgcc aatgatgtta cagatgagat 6360 ggtcagacta aactggctga cggaatttat gcctcttccg accatcaagc attttatccg 6420 tactcctgat gatgcatggt tactcaccac tgcgatcccc gggaaaacag cattccaggt 6480 attagaagaa tatcctgatt caggtgaaaa tattgttgat gcgctggcag tgttcctgcg 6540 ccggttgcat tcgattcctg tttgtaattg tccttttaac agcgatcgcg tatttcgtct 6600 cgctcaggcg caatcacgaa tgaataacgg tttggttgat gcgagtgatt ttgatgacga 6660 gcgtaatggc tggcctgttg aacaagtctg gaaagaaatg cataagcttt tgccattctc 6720 accggattca gtcgtcactc atggtgattt ctcacttgat aaccttattt ttgacgaggg 6780 gaaattaata ggttgtattg atgttggacg agtcggaatc gcagaccgat accaggatct 6840 tgccatccta tggaactgcc tcggtgagtt ttctccttca ttacagaaac ggctttttca 6900 aaaatatggt attgataatc ctgatatgaa taaattgcag tttcatttga tgctcgatga 6960 gtttttctaa tcagaattgg ttaattggtt gtaacactgg cagagcatta cgctgacttg 7020 acgggacggc ggctttgttg aataaatcga acttttgctg agttgaagga tcagatcacg 7080 tccacctaca acaaagctct catcaaccgt ggctccctca ctttctggct ggatgatggg 7200 gcgattcagg cctggtatga gtcagcaaca ccttcttcac gaggcagacc tcagcgcccc 7260 ccccccctg caggtcccac ggcggcggtg ctcaacggcc tcaacctact actgggctgc 7320 ttcctaatgc aggagtcgca taagggagag cgtcgagtat ctatgattgg aagtatggga 7380 atggtgatac ccgcattctt cagtgtcttg aggtctccta tcagattatg cccaactaaa 7440 gcaaccggag gaggagattt catggtaaat ttctctgact tttggtcatc agtagactcg 7500 aactgtgaga ctatctcggt tatgacagca gaaatgtcct tcttggagac agtaaatgaa 7560 gtcccaccaa taaagaaatc cttgttatca ggaacaaact tcttgtttcg aactttttcg 7620 gtgccttgaa ctataaaatg tagagtggat atgtcgggta ggaatggagc gggcaaatgc 7680 ttaccttctg gaccttcaag aggtatgtag ggtttgtaga tactgatgcc aacttcagtg 7740 acaacgttgc tatttcgttc aaaccattcc gaatccagag aaatcaaagt tgtttgtcta 7800 ctattgatcc aagccagtgc ggtcttgaaa ctgacaatag tgtgctcgtg ttttgaggtc 7860 atctttgtat gaataaatct agtctttgat ctaaataatc ttgacgagcc aaggcgataa 7920 atacccaaat ctaaaactct tttaaaacgt taaaaggaca agtatgtctg cctgtattaa 7980 accccaaatc agctcgtagt ctgatcctca tcaacttgag gggcactatc ttgttttaga 8040 gaaatttgcg gagatgcgat atcgagaaaa aggtacgctg attttaaacg tgaaatttat 8100 ctcaagatct ctgcctcgcg cgtttcggtg atgacggtga aaacctctga cacatgcagc 8160 tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg 8220 gcgcgtcagc gggtgttggc gggtgtcggg gcgcagccat gacccagtca cgtagcgata 8280 gcggagtgta tactggctta actatgcggc atcagagcag attgtactga gagtgcacca 8340 tatgcggtgt gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggcgctcttc 8400 cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc 8460 tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaacat 8520 gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt 8580 ccataggete egececetg acgageatea caaaaatega egeteaagte agaggtggeg 8640 aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc 8700 tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt 8760 ggcgctttct caatgctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa 8820 gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta 8880 tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa 8940 caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa 9000 ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt 9060 cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt 9120 ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat 9180 cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat 9240 gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc 9300 aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc 9360 acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta 9420 gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga 9480 cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg 9540 cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc 9600

tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctgcaggcat 9660 cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag 9720 gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat 9780 cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa 9840 ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa 9900 gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt caacacggga 9960 taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg 10020 gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc 10080 acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg 10140 aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact 10200 cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat 10260 atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 10320 gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa ataggcgtat 10380 cacgaggccc tttcgtcttc aagaattaat tctcatgttt gacagcttat catcgataag 10440 ctgactcatg ttggtattgt gaaatagacg cagatcggga acactgaaaa ataacagtta 10500 10512 ttattcgaga tc

<210> 34 <211> 10462 <212> DNA

<213> Plasmid pUB116

<400> 34

gatctaacat ccaaagacga aaggttgaat gaaacctttt tgccatccga catccacagg 60 tccattctca cacataagtg ccaaacgcaa caggagggga tacactagca gcagaccgtt 120 gcaaacgcag gacctccact cctcttctcc tcaacaccca cttttgccat cgaaaaacca 180 gcccagttat tgggcttgat tggagctcgc tcattccaat tccttctatt aggctactaa 240 caccatgact ttattagcct gtctatcctg gcccccctgg cgaggttcat gtttgtttat 300 ttccgaatgc aacaagctcc gcattacacc cgaacatcac tccagatgag ggctttctga 360 gtgtggggtc aaatagtttc atgttcccca aatggcccaa aactgacagt ttaaacgctg 420 tcttggaacc taatatgaca aaagcgtgat ctcatccaag atgaactaag tttggttcgt 480 tgaaatgeta aeggeeagtt ggteaaaaag aaactteeaa aagtegeeat aeegtttgte 540 ttgtttggta ttgattgacg aatgctcaaa aataatctca ttaatgctta gcgcagtctc 600 tctatcgctt ctgaaccccg gtgcacctgt gccgaaacgc aaatggggaa acacccgctt 660 tttggatgat tatgcattgt ctccacattg tatgcttcca agattctggt gggaatactg 720 ctgatagcct aacgttcatg atcaaaattt aactgttcta acccctactt gacagcaata 780 tataaacaga aggaagctgc cctgtcttaa accttttttt ttatcatcat tattagctta 840 ctttcataat tgcgactggt tccaattgac aagcttttga ttttaacgac ttttaacgac 900 aacttgagaa gatcaaaaaa caactaatta ttcgaaggat ccaaacgatg agatttcctt 960 caatttttac tgcagtttta ttcgcagcat cctccgcatt agctgctcca gtcaacacta 1020 caacagaaga tgaaacggca caaattccgg ctgaagctgt catcggttac tcagatttag 1080 aaggggattt cgatgttgct gttttgccat tttccaacag cacaaataac gggttattgt 1140 ttataaatac tactattgcc agcattgctg ctaaagaaga aggggtatct ctcgagaaaa 1200 gatgcggacc gggcaggggg ttcgggaaga ggaggcaccc caaaaagctg acccctttag 1260 cctacaagca gtttatcccc aatgtggccg agaagaccct aggcgccagc ggaaggtatg 1320 aagggaagat ctccagaaac tccgagcgat ttaaggaact caccccaat tacaaccccg 1380 acatcatatt taaggatgaa gaaaacaccg gagcggacag gctgatgact cagaggtgta 1440 aggacaagtt gaacgctttg gccatctcgg tgatgaacca gtggccagga gtgaaactgc 1500 gggtgaccga gggctgggac gaagatggcc accactcaga ggagtctctg cactacgagg 1560 gccgcgcagt ggacatcacc acgtctgacc gcgaccgcag caagtacggc atgctggccc 1620 gcctggcggt ggaggccggc ttcgactggg tgtactacga gtccaaggca catatccact 1680 gctcggtgaa agcagagaac tcggtggcgg ccaaatcggg aggcgtcgac aaaactcaca 1740 catgcccacc gtgcccagca cctgaactcc tggggggacc gtcagtcttc ctcttccccc 1800 caaaacccaa ggacaccctc atgatctccc ggacccctga ggtcacatgc gtggtggtgg 1860 acgtgagcca cgaagaccct gaggtcaagt tcaactggta cgtggacggc gtggaggtgc 1920 ataatgccaa gacaaagccg cgggaggagc agtaccagag cacgtaccgt gtggtcagcg 1980

```
tcctcaccgt cctgcaccag gactggctga atggcaagga gtacaagtgc aaggtctcca 2040
acaaagccct cccagccccc atcgagaaaa ccatctccaa agccaaaggg cagccccgag 2100
aaccacaggt gtacaccctg cccccatccc gggatgagct gaccaagaac caggtcagcc 2160
tgacctgcct ggtcaaaggc ttctatccca gcgacatcgc cgtggagtgg gagagcaatg 2220
ggcagccgga gaacaactac aagaccacgc ctcccgtgtt ggactccgac ggctccttct 2280
tectetacag caageteace gtggacaaga geaggtggea geaggggaae gtetteteat 2340
geteegtgat geatgagget etgeacaace actacaegea gaagageete teeetgtete 2400
ccgggaaatg agtgcggcgg ccgcgaatta attcgcctta gacatgactg ttcctcagtt 2460
caagttgggc acttacgaga agaccggtct tgctagattc taatcaagag gatgtcagaa 2520
tgccatttgc ctgagagatg caggcttcat ttttgatact tttttatttg taacctatat 2580
agtataggat tttttttgtc attttgtttc ttctcgtacg agcttgctcc tgatcagcct 2640
atctcgcagc tgatgaatat cttgtggtag gggtttggga aaatcattcg agtttgatgt 2700
ttttcttggt atttcccact cctcttcaga gtacagaaga ttaagtgaga agttcgtttg 2760
tgcaagctta tcgataagct ttaatgcggt agtttatcac agttaaattg ctaacgcagt 2820
caggcaccgt gtatgaaatc taacaatgcg ctcatcgtca tcctcggcac cgtcaccctg 2880
gatgctgtag gcataggctt ggttatgccg gtactgccgg gcctcttgcg ggatatcgtc 2940
catteegaca geategeeag teactatgge gtgetgetag egetatatge gttgatgeaa 3000
tttctatgcg cacccgttct cggagcactg tccgaccgct ttggccgccg cccagtcctg 3060
ctcgcttcgc tacttggagc cactatcgac tacgcgatca tggcgaccac acccgtcctg 3120
tggatctatc gaatctaaat gtaagttaaa atctctaaat aattaaataa gtcccagttt 3180
ctccatacga accttaacag cattgcggtg agcatctaga ccttcaacag cagccagatc 3240
catcactgct tggccaatat gtttcagtcc ctcaggagtt acgtcttgtg aagtgatgaa 3300
cttctggaag gttgcagtgt taactccgct gtattgacgg gcatatccgt acgttggcaa 3360
agtgtggttg gtaccggagg agtaatctcc acaactctct ggagagtagg caccaacaaa 3420
cacagatcca gcgtgttgta cttgatcaac ataagaagaa gcattctcga tttgcaggat 3480
caagtgttca ggagcgtact gattggacat ttccaaagcc tgctcgtagg ttgcaaccga 3540
tagggttgta gagtgtgcaa tacacttgcg tacaatttca accettggca actgcacage 3600
ttggttgtga acagcatctt caattctggc aagctccttg tctgtcatat cgacagccaa 3660
cagaatcacc tgggaatcaa taccatgttc agcttgagca gaaggtctga ggcaacgaaa 3720
tctggatcag cgtatttatc agcaataact agaacttcag aaggcccagc aggcatgtca 3780
atactacaca gggctgatgt gtcattttga accatcatct tggcagcagt aacgaactgg 3840
tttcctggac caaatatttt gtcacactta ggaacagttt ctgttccgta agccatagca 3900
gctactgcct gggcgcctcc tgctagcacg atacacttag caccaacctt gtgggcaacg 3960
tagatgactt ctggggtaag ggtaccatcc ttcttaggtg gagatgcaaa aacaatttct 4020
ttgcaaccag caactttggc aggaacaccc agcatcaggg aagtggaagg cagaattgcg 4080
gttccaccag gaatatagag gccaactttc tcaataggtc ttgcaaaacg agagcagact 4140
acaccaggge aagteteaac ttgeaacgte teegttagtt gagetteatg gaattteetg 4200
acgttatcta tagagagatc aatggctctc ttaacgttat ctggcaattg cataagttcc 4260
tctgggaaag gagcttctaa cacaggtgtc ttcaaagcga ctccatcaaa cttggcagtt 4320
agttctaaaa gggctttgtc accattttga cgaacattgt cgacaattgg tttgactaat 4380
tccataatct gttccgtttt ctggatagga cgacgaaggg catcttcaat ttcttgtgag 4440
gaggccttag aaacgtcaat tttgcacaat tcaatacgac cttcagaagg gacttcttta 4500
ctagtgacct ttagggactt catatccagg tttctctcca cctcgtccaa cgtcacaccg 4620
tacttggcac atctaactaa tgcaaaataa aataagtcag cacattccca ggctatatct 4680
teettggatt tagettetge aagtteatea getteeteee taattttage gtteaaacaa 4740
aacttcgtcg tcaaataacc gtttggtata agaaccttct ggagcattgc tcttacgatc 4800
ccacaaggtg cttccatggc tctaagaccc tttgattggc caaaacagga agtgcgttcc 4860
aagtgacaga aaccaacacc tgtttgttca accacaaatt tcaagcagtc tccatcacaa 4920
tecaattega taeceageaa ettttgagtt egteeagatg tageacettt ataecacaaa 4980
ccgtgacgac gagattggta gactccagtt tgtgtcctta tagcctccgg aatagacttt 5040
ttggacgagt acaccaggcc caacgagtaa ttagaagagt cagccaccaa agtagtgaat 5100
agaccatcgg ggcggtcagt agtcaaagac gccaacaaaa tttcactgac agggaacttt 5160
ttgacatctt cagaaagttc gtattcagta gtcaattgcc gagcatcaat aatggggatt 5220
ataccagaag caacagtgga agtcacatct accaactttg cggtctcaga aaaagcataa 5280
acagttctac taccgccatt agtgaaactt ttcaaatcgc ccagtggaga agaaaaaggc 5340
acagcgatac tagcattagc gggcaaggat gcaactttat caaccagggt cctatagata 5400
```

accetagege etgggateat cetttggaea actetttetg ceaaatetag gteeaaate 5460 acttcattga taccattatt gtacaacttg agcaagttgt cgatcagctc ctcaaattgg 5520 tcctctgtaa cggatgactc aacttgcaca ttaacttgaa gctcagtcga ttgagtgaac 5580 ttgatcaggt tgtgcagctg gtcagcagca tagggaaaca cggcttttcc taccaaactc 5640 aaggaattat caaactctgc aacacttgcg tatgcaggta gcaagggaaa tgtcatactt 5700 gaagtcggac agtgagtgta gtcttgagaa attctgaagc cgtattttta ttatcagtga 5760 gtcagtcatc aggagatcct ctacgccgga cgcatcgtgg ccgacctgca ggtcggcatc 5820 accggcgcca caggtgcggt tgctggcgcc tatatcgccg acatcaccga tggggaagat 5880 cgggctcgcc acttcgggct catgagcgct tgtttcggcg tgggtatggt ggcaggcccc 5940 gtggccgggg gactgttggg cgccatctcc ttggacctgc aggggggggg ggggaaagcc 6000 acgttgtgtc tcaaaatctc tgatgttaca ttgcacaaga taaaaatata tcatcatgaa 6060 caataaaact gtctgcttac ataaacagta atacaagggg tgttatgagc catattcaac 6120 gggaaacgtc ttgctcaagg ccgcgattaa attccaacat ggatgctgat ttatatgggt 6180 ataaatgggc tcgcgataat gtcgggcaat caggtgcgac aatctatcga ttgtatggga 6240 agcccgatgc gccagagttg tttctgaaac atggcaaagg tagcgttgcc aatgatgtta 6300 cagatgagat ggtcagacta aactggctga cggaatttat gcctcttccg accatcaagc 6360 attttatccg tactcctgat gatgcatggt tactcaccac tgcgatcccc gggaaaacag 6420 cattccaggt attagaagaa tatcctgatt caggtgaaaa tattgttgat gcgctggcag 6480 tgttcctgcg ccggttgcat tcgattcctg tttgtaattg tccttttaac agcgatcgcg 6540 tatttcgtct cgctcaggcg caatcacgaa tgaataacgg tttggttgat gcgagtgatt 6600 ttgatgacga gcgtaatggc tggcctgttg aacaagtctg gaaagaaatg cataagcttt 6660 tgccattctc accggattca gtcgtcactc atggtgattt ctcacttgat aaccttattt 6720 ttgacgaggg gaaattaata ggttgtattg atgttggacg agtcggaatc gcagaccgat 6780 accaggatet tgecateeta tggaactgee teggtgagtt tteteettea ttacagaaac 6840 ggctttttca aaaatatggt attgataatc ctgatatgaa taaattgcag tttcatttga 6900 tgctcgatga gtttttctaa tcagaattgg ttaattggtt gtaacactgg cagagcatta 6960 cgctgacttg acgggacggc ggctttgttg aataaatcga acttttgctg agttgaagga 7020 caccaactgg tccacctaca acaaagctct catcaaccgt ggctccctca ctttctggct 7140 ggatgatggg gcgattcagg cctggtatga gtcagcaaca ccttcttcac gaggcagacc 7200 tragegere cerecetty raggiterear ggeggeggty etraacgger traacctact 7260 actgggctgc ttcctaatgc aggagtcgca taagggagag cgtcgagtat ctatgattgg 7320 aagtatggga atggtgatac ccgcattctt cagtgtcttg aggtctccta tcagattatg 7380 cccaactaaa gcaaccggag gaggagattt catggtaaat ttctctgact tttggtcatc 7440 agtagactcg aactgtgaga ctatctcggt tatgacagca gaaatgtcct tcttggagac 7500 agtaaatgaa gtcccaccaa taaagaaatc cttgttatca ggaacaaact tcttgtttcg 7560 aactttttcg gtgccttgaa ctataaaatg tagagtggat atgtcgggta ggaatggagc 7620 gggcaaatgc ttaccttctg gaccttcaag aggtatgtag ggtttgtaga tactgatgcc 7680 aacttcagtg acaacgttgc tatttcgttc aaaccattcc gaatccagag aaatcaaagt 7740 tgtttgtcta ctattgatcc aagccagtgc ggtcttgaaa ctgacaatag tgtgctcgtg 7800 ttttgaggtc atctttgtat gaataaatct agtctttgat ctaaataatc ttgacgagcc 7860 aaggcgataa atacccaaat ctaaaactct tttaaaacgt taaaaggaca agtatgtctg 7920 cctgtattaa accccaaatc agctcgtagt ctgatcctca tcaacttgag gggcactatc 7980 ttgttttaga gaaatttgcg gagatgcgat atcgagaaaa aggtacgctg attttaaacg 8040 tgaaatttat ctcaagatct ctgcctcgcg cgtttcggtg atgacggtga aaacctctga 8100 cacatgcagc tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa 8160 gcccgtcagg gcgcgtcagc gggtgttggc gggtgtcggg gcgcagccat gacccagtca 8220 cgtagcgata gcggagtgta tactggctta actatgcggc atcagagcag attgtactga 8280 gagtgcacca tatgcggtgt gaaataccgc acagatgcgt aaggagaaaa taccgcatca 8340 ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag 8400 cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag 8460 gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc 8520 tggcgttttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc 8580 agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc 8640 tegtgegete teetgtteeg accetgeege ttaceggata cetgteegee ttteteeett 8700 cgggaagcgt ggcgctttct caatgctcac gctgtaggta tctcagttcg gtgtaggtcg 8760 ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat 8820

ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag 8880 ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt 8940 ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc 9000 cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta 9060 gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag 9120 atcctttgat cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga 9180 ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa 9240 gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa 9300 tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc 9360 ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga 9420 taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa 9480 gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt 9540 gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg 9600 ctgcaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc 9660 aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg 9720 gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag 9780 cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt 9840 actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt 9900 caacacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac 9960 gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac 10020 ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag 10080 caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa 10140 tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga 10200 gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc 10260 cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa 10320 ataggcgtat cacgaggccc tttcgtcttc aagaattaat tctcatgttt gacagcttat 10380 catcgataag ctgactcatg ttggtattgt gaaatagacg cagatcggga acactgaaaa 10440 10462 ataacagtta ttattcgaga tc

```
<210> 35
<211> 4205
<212> DNA
<213> Plasmid pEAG657
```

<400> 35

ctaaattgta agcgttaata ttttgttaaa attcgcgtta aatttttgtt aaatcagctc 60 attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga 120 gatagggttg agtgttgttc cagtttggaa caagagtcca ctattaaaga acgtggactc 180 caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc 240 ctaatcaagt tttttggggt cgaggtgccg taaagcacta aatcggaacc ctaaagggag 300 cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa 360 agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac 420 cacaccegee gegettaatg egeegetaea gggegegtee cattegeeat teaggetgeg 480 caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg 540 gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg 600 taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tgggtaccgg 660 gecetetaga teettteage teeetgeece ggacatgeec agtgggtgga agetgeecte 720 ttctagcagg agacgcccca ggcggtagag cagctggggg taccaatgca caccctcccc 780 cggagtccag ctgccccatg ccaagctgtg aaagagtctc aggggccaga aggccaactg 840 agccaggtgg tggtcagcca cggccgcgaa gcaggatgcc accacatcct ccaccaccag 900 tgtcccatgc tttgtgagcg gggcgtaggc cccgagggcc acgtgtgtag agacagctgc 960 cacgegggea ggetgeagge etggeacece agecaceage aegtaetgge eaggetgeae 1020 gtggctggca aatgtggccc ggaagcgggc tgccggctcc gtgtgattgt cagccgtaaa 1080 gagcaggtga gcgggtgtga gtgccaggcg gcgtgggggg tcctgagtct cgatgacctg 1140 gaaggetete ageetgtggg getegeggte caggaaaatg ageacatege tgaaggtggg 1200 gctcccatcc tcccccatgg ccagcacacg gtctcccggc ctcacggctg acaaggccac 1260

acgcgcccca ctctccaggc gtacctgggc tgcggccgcg aatcagccgc ccgtcttggc 1320 tgcggccgag tgctcggact tgacggagca atgcacgtgg gcctttgact cgtaatacac 1380 ccagtcaaag ccggcctcca ctgccaagcg cgccagcagt ccatacttat tgcggtcgcg 1440 gtctgatgtg gtgatgtcca ccgcgcggcc ctcataatgc agggactcct ctgagtggtg 1500 gccgtcctcg tcccagccct cggtcacccg cagcttcaca ccgggccact ggttcatcac 1560 cgagatagcc agcgagttca ggcggtcctt gcagcgctgg gtcatgaggc ggtcggcgc 1620 tgtgttctcc tcgtccttga agatgatgtc tggattgtaa ttggggggtga gctccttgaa 1680 gegeteggag etgegagega tettgeette atagegteeg etggegeeea gggtettete 1740 gggcacattg gggctgaact gcttgtaggc gagcggcacg agtttgcgtg gcggtcgccg 1800 geggetgeee accaecegae eeggeeegea geeecatgee geeggeaeea eeageageag 1860 caacaggacc aggcagaagt gcagtcgggg ccggagccgg gcgggagaca tggcggccgc 1920 gacggtatcg ataagcttga tatcgaattc ctgcagcccg ggggatccac tagttctaga 1980 geggeegeea eegeggtgga geteeagett ttgtteeett tagtgagggt taattgegeg 2040 cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc 2100 acacaacata cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta 2160 actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca 2220 gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc 2280 cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc 2340 tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaacat 2400 gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt 2460 ccataggete egececetg acgageatea caaaaatega egeteaagte agaggtggeg 2520 aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc 2580 tectgtteeg accetgeege ttaceggata cetgteegee ttteteeett egggaagegt 2640 ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa 2700 gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta 2760 tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa 2820 caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa 2880 ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt 2940 cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt 3000 ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat 3060 cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat 3120 gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc 3180 aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc 3240 acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta 3300 gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga 3360 cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg 3420 cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc 3480 tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat 3540 cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag 3600 gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat 3660 cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa 3720 ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgacg cgtcaaccaa 3780 gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt caatacggga 3840 taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg 3900 gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc 3960 acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg 4020 aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact 4080 cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat 4140 atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 4200 gccac 4205

<210> 36 <211> 30

<212> DNA

<213> Artificial Sequence

<220> <223> Description of Artificial Sequence: Oligonucleotide	
<400> 36	
tcgagaaaag atgcggaccg ggcagggggt	30
<210> 37	
<211> 29	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence:	
Oligonucleotide	
<400> 37	
cgaaccccct gcccggtccg catctttc	29
<210 > 38	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence: Primer	
(223) Description of Artificial Sequence. Illiner	
<400> 38	
tcaggatgca tttgacagtg actgg	25
<210> 39	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence: Primer	
<400> 39	
actecgagte ggaggaatea gacee	25
accegagee ggaggaacea gaeee	23
<210> 40	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence: Primer	
<400> 40	0.0
cgaagtggtg aagttcatgg atg	23

```
<210> 41
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Primer
<400> 41
ttctgtatca gtctttcctg gtgag
                                                                    25
<210> 42
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Primer
<400> 42
tacaacttca agcagaagag
                                                                    20
<210> 43
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Primer
<400> 43
cagctcttag cagacattgg
                                                                    20
<210> 44
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Primer
<400> 44
                                                                    24
caacacaac gctctgcaga gaga
<210> 45
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Primer
<400> 45
```

gtctgatgtg gtgatgtcca ccgcgcggcc ctcataatgc agggactcct ctgagtggtg 1500 gccgtcctcg tcccagccct cggtcacccg cagcttcaca ccgggccact ggttcatcac 1560 cgagatagcc agcgagttca ggcggtcctt gcagcgctgg gtcatgaggc ggtcggcc 1620 tgtgttctcc tcgtccttga agatgatgtc tggattgtaa ttgggggtga gctccttgaa 1680 gcgctcggag ctgcgagcga tcttgccttc atagcgtccg ctggcgccca gggtcttctc 1740 gggcacattg gggctgaact gcttgtaggc gagcggcacg agtttgcgtg gcggtcgccg 1800 geggetgeec accaceegae eeggeeegea geeceatgee geeggeaeea eeageageag 1860 caacaggacc aggcagaagt gcagtcgggg ccggagccgg gcgggagaca tggcggccgc 1920 gacggtatcg ataagcttga tatcgaattc ctgcagcccg ggggatccac tagttctaga 1980 gcggccgcca ccgcggtgga gctccagctt ttgttccctt tagtgagggt taattgcgcg 2040 cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc 2100 acacaacata cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta 2160 actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca 2220 gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc 2280 cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc 2340 tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaacat 2400 gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt 2460 ccataggete egececetg acgageatea caaaaatega egeteaagte agaggtggeg 2520 aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc 2580 tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt 2640 ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa 2700 gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta 2760 tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa 2820 caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa 2880 ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt 2940 cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt 3000 ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat 3060 cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat 3120 gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc 3180 aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc 3240 acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta 3300 gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga 3360 cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg 3420 cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc 3480 tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat 3540 cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag 3600 gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat 3660 cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa 3720 ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgacg cgtcaaccaa 3780 gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt caatacggga 3840 taataccqcq ccacataqca gaactttaaa agtgctcatc attggaaaac gttcttcggg 3900 gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc 3960 acccaactga tetteageat ettttaettt caccagegtt tetgggtgag caaaaacagg 4020 aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact 4080 cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat 4140 atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 4200 4205 gccac